CRESCIMENTO ECONÔMICO NO ESTADO DE SÃO PAULO

UMA ANÁLISE ESPACIAL

RODRIGO DE SOUZA VIEIRA

Crescimento econômico no estado de São Paulo

RODRIGO DE SOUZA VIEIRA

CRESCIMENTO ECONÔMICO NO ESTADO DE SÃO PAULO

UMA ANÁLISE ESPACIAL

CULTURA ACADÊMICA
Editora

Cultura Acadêmica
Praça da Sé, 108
01001-900 – São Paulo – SP
Tel.: (0xx11) 3242-7171
Fax: (0xx11) 3242-7172
www.editoraunesp.com.br
feu@editora.unesp.br

CIP – Brasil. Catalogação na fonte
Sindicato Nacional dos Editores de Livros, RJ

V715c

Vieira, Rodrigo de Souza
Crescimento econômico no estado de São Paulo : uma análise espacial / Rodrigo de Souza Vieira. – São Paulo : Cultura Acadêmica, 2009.

Inclui bibliografia
ISBN 978-85-7983-013-6

1. Desenvolvimento econômico – São Paulo (Estado). 2. São Paulo (Estado) – Condições econômicas. 3. São Paulo (Estado) – Condições sociais. 4. Econometria. I. Título.

09-6212. CDD: 338.98161
CDU: 338.1(815.61)

Este livro é publicado pelo Programa de Publicações Digitais da Pró-Reitoria de Pós-Graduação da Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP)

Editora afiliada:

Asociación de Editoriales Universitarias de América Latina y el Caribe

Associação Brasileira de Editoras Universitárias

Agradeço a Deus, a minha família e aos amigos, que forneceram a base para ultrapassar mais essa etapa da minha vida. Agradeço aos amigos de Viçosa, ao pessoal de Araraquara pelos dois anos que ficarão marcados para sempre e aos amigos de São Paulo. Aos professores da pós-graduação, em especial ao professor Alexandre Sartoris, e aos professores Danilo Igliori e Sara Prado, que foram decisivos para o sucesso deste trabalho.

Sumário

Introdução

Na atualidade, São Paulo é o estado economicamente mais importante do País, pois responde por algo em torno de 34% do PIB nacional, com uma população que representa aproximados 22% da população total brasileira. Além disso, o Estado detém parcela significativa da indústria tecnologicamente mais avançada e boa parte da mão de obra qualificada do país. Entretanto, a despeito do tamanho de sua população e de toda a grandeza de seu PIB em relação aos demais estados, talvez a economia paulista não tenha sido convenientemente estudada em sua complexidade espacial e geográfica.

Sob o ponto de vista populacional, São Paulo tem uma população comparável à da Argentina.[1] Sua capital é o centro de uma aglomeração urbana que faz da mesma uma das maiores cidades do mundo, com uma população absoluta de aproximadamente 11 milhões de pessoas, além de uma densidade populacional de 7.175 pessoas por km^2. Algumas das cidades que compõem a Região Metropolitana de São Paulo (RMSP) estão entre as maiores do País e têm importância econômica indiscutível, como as cidades do ABC (Santo André, São Bernardo do Campo e São Caetano do Sul) e Guarulhos. Em termos

---

1 Estimativas da Fundação Seade apontam para uma população em torno de 40 milhões de habitantes em 2007.

populacionais, a região possui em torno de 48% da população total do estado.

Sob o ponto de vista econômico, a RMSP responde por mais da metade do PIB estadual, sendo que, no ano de 2004, sua participação era de 50,3%. Além da Grande São Paulo, há outras duas regiões metropolitanas, a de Campinas e a da Baixada Santista[2] que, juntamente com São José dos Campos e Sorocaba, formam o entorno da RMSP e delimitam a área de maior desenvolvimento econômico do Estado, respondendo por cerca de 83% do PIB estadual.

A atividade econômica, entretanto, não se restringe à região metropolitana e seu entorno. A região central do estado também é um polo econômico importante, na qual se destacam as cidades de Ribeirão Preto, especialmente por meio do setor comercial, e São Carlos, importante centro tecnológico. No oeste do estado, cidades como Presidente Prudente e São José do Rio Preto possuem economias com alto grau de desenvolvimento e dinamismo, e destacam-se pelo elevado padrão de vida da população.

Contudo, paralelo a economias fortalecidas e com alto nível de produção e renda, a economia paulista apresenta regiões pobres, como o Vale do Ribeira, além dos bolsões de pobreza situados em diversos locais, destacando-se, neste aspecto, a própria Região Metropolitana de São Paulo.[3]

As diferenças de dinamismo também se verificam através do estado. Percebe-se que as experiências de crescimento dos municípios paulistas têm variado amplamente, uma vez que a população de

---

2 São três as regiões metropolitanas do Estado, a saber: (1) Região Metropolitana de São Paulo, criada em 8/6/1973 pela Lei Complementar (LC) Federal 14/73, que abrange 39 municípios; (2) Região Metropolitana da Baixada Santista, criada em 30/7/1996 pela LC Estadual 815/96, compondo-se de 9 municípios; e (3) Região Metropolitana de Campinas, criada em 19/6/2000 pela LC Estadual 870/00, que abrange 19 municípios.

3 Sobre este ponto, Cano (2002, p. 284) afirma: "Em que pese a região metropolitana de São Paulo ter tido, em 2000, uma renda média por habitante em torno de US$5.000,00 (68% acima da média nacional), ali se encontravam 5,2 milhões de pobres (ou 30% de sua população), perfazendo 10% do número de pobres do país".

algumas cidades cresceu vertiginosamente enquanto outras enfrentaram queda em sua população. Nos últimos anos, segundo dados da Fundação Seade, as cidades paulistas têm apresentado taxas médias de crescimento maiores do que o restante do País, o que reflete o forte poder de atração que o estado tem em relação aos demais estados da Federação. Tal comportamento pode ser atribuído à concentração das atividades produtivas e sua capacidade de geração de renda. No entanto, diversos municípios vêm apresentando taxas negativas, com queda contínua de sua população, sendo que a maior parte deles se concentra nas regiões oeste e sul do estado.

Nesse sentido, este trabalho busca o entendimento para a seguinte questão: por que algumas cidades do estado foram mais bem-sucedidas do que outras nos últimos anos? A prosperidade das cidades paulistas é resultado de fatores externos tais como localização ou choques setoriais? Ou então, resultado de políticas públicas individuais empreendidas por seus governantes? A compreensão da participação de forças externas e de esforços internos de políticas nesse processo se faz importante para desvendar o alcance potencial que políticas intervencionistas possam vir a ter.

No caso específico de São Paulo, parece haver uma relação direta entre o comportamento da ocupação territorial e a localização das atividades industriais.[4] Segundo estudos empíricos, entre eles Diniz & Crocco (1996), Cano (2002) e Diniz (2002), o processo de desconcentração industrial verificado principalmente a partir da década de 1970, que alterou de modo significativo a configuração regional da produção do Estado, favoreceu cidades fora da RMSP e provocou uma redistribuição da população. De fato, tal processo não pode ser relegado a segundo plano quando se trata de estudar espacialmente o crescimento econômico em São Paulo.

Ademais, depreende-se da literatura de crescimento econômico que fatores como nível de renda inicial (Solow, 1956), nível educacional da população (Lucas 1988, Mankiw, Romer & Weil, 1992) e infraestrutura social (Barro, 1990) são responsáveis pelo comporta-

4 Atlas Seade da economia paulista.

mento das taxas de crescimento dos países. Recentemente, modelos vêm sendo criados no sentido de utilizar a estrutura teórica desenvolvida para países no estudo de regiões – um exemplo é o trabalho de Barro & Sala-i-Martin (1995). Nesse sentido, este trabalho utiliza as considerações dessa nova corrente no intuito de identificar empiricamente quais os fatores que determinam o crescimento econômico dos municípios paulistas.

Além disso, busca-se contribuir com a literatura ao inserir a questão espacial como crucial para o entendimento a respeito de quais fatores influenciam o crescimento das regiões. Considera-se, dessa forma, a importância das externalidades geográficas como fator determinante de retornos adicionais, advindos da aglomeração de firmas e pessoas (trabalhadores) em uma determinada localidade.

Para tratar as questões relativas à localização, utilizam-se, como referência, os trabalhos da Nova Geografia Econômica (NGE). Segundo a NGE, atribui-se a variáveis adicionais a responsabilidade pelo desempenho econômico das regiões. Destacam-se variáveis como densidade populacional (Fujita et al., 1999; Fujita & Thisse, 2002), taxa de urbanização (Fujita et al., 1999; Fujita & Thisse, 2002), desigualdade interpessoal da renda (Alesina & Rodrick, 1994) e taxa de participação do emprego industrial (Fingleton, 1999), que consistem em determinantes do comportamento regional com relação à produtividade e à qualidade de vida.

Em linhas gerais, este estudo busca comparar o crescimento dos municípios paulistas por meio de fatores que o expliquem, levando-se em conta externalidades geográficas. Mais especificamente, o trabalho busca: (1) verificar quais variáveis são correlacionadas com as taxas de crescimento dos municípios paulistas, (2) identificar o tipo de influência das externalidades espaciais na trajetória de crescimento desses municípios, captando seus efeitos, e (3) identificar o tipo de interação espacial que melhor descreve o padrão apresentado pelos dados, a fim de contribuir para a discussão sobre as diferentes matrizes de pesos espaciais utilizadas na literatura de econometria espacial.

O último objetivo pauta-se na discussão referente à utilização da matriz de pesos espaciais com o intuito de identificar possíveis efeitos

de transbordamento entre as regiões. A principal diferença entre a econometria tradicional e a teoria econométrica espacial situa-se na utilização, por parte desta, de uma medida de ponderação que capta uma possível influência entre as variáveis de unidades contíguas à unidade em estudo. Dessa forma, a econometria espacial admite que uma regressão possa apresentar erros espacialmente correlacionados. Essa medida de ponderação consiste, justamente, na matriz de pesos espaciais, consensualmente denominada "matriz W". Entretanto, a literatura de econometria espacial admite que a escolha da matriz de pesos permite uma certa arbitrariedade por parte do pesquisador. Quando a matriz de pesos é construída, é tratada como um fator exógeno, uma vez que é determinada *a priori*. O pesquisador pressupõe, de antemão, uma estrutura específica para os erros do modelo.

Dada a natureza *ad hoc* da escolha da matriz W, este trabalho procura avançar na discussão a respeito, adotando, como pano de fundo, os dados referentes aos municípios paulistas.

Em suma, o trabalho tem como objetivo, a princípio, identificar os determinantes do crescimento econômico no estado de São Paulo, controlando para possíveis influências espaciais, e, em um segundo momento, contribuir com a literatura de econometria espacial no sentido de testar diversos tipos de matrizes de pesos, tentando, com isso, encontrar a matriz W mais adequada para a estrutura de correlação espacial do modelo considerado.

As principais contribuições deste trabalho consistem em: (1) revisão da literatura pertinente ao assunto e seu ordenamento sistemático, (2) teste empírico para os municípios paulistas do modelo de crescimento proposto por Glaeser et al. (1995), com o acréscimo de parâmetros espaciais, e (3) discussão a respeito da matriz de pesos espaciais mais adequada para a amostra de dados levantada.

1
REVISÃO DE LITERATURA

O crescimento econômico das nações

Os estudos a respeito dos determinantes do crescimento econômico das cidades e regiões estiveram, de forma geral, ligados à grande teoria de crescimento econômico das nações, principalmente, aqueles balizados pela literatura econômica *mainstrean*. Barro & Sala-i-Martin (1995) discutem os principais conceitos e formulações teóricas sobre crescimento econômico sugeridos no decorrer do século XX, e utilizam as ferramentas teóricas propostas em análises de âmbito regional (regiões europeias), estadual (estados norte-americanos) e municipal (municípios japoneses). Os autores apontam para as similaridades analíticas observadas no comportamento das distintas unidades geográficas.

Para Barro & Sala-i-Martin, o ponto de partida da moderna teoria do crescimento econômico é o artigo clássico de Ramsey (1928), o qual, para os referidos autores, consistiu em um trabalho várias décadas à frente de seu tempo. Nos anos 50, a teoria de crescimento econômico ganhou dimensão com os trabalhos de Solow (1956) e Swan (1956), que se valeram de ingredientes fornecidos por economistas clássicos, tais como: Adam Smith (1776), David Ricardo (1817), Thomas Malthus (1798) e economistas "não tão clássicos",

como o próprio Ramsey (1928), Allyn Young (1928), Frank Knight (1944) e Joseph Schumpeter (1934), para construir seus modelos de interpretação dos determinantes do crescimento econômico de longo prazo das nações.

O modelo "Solow-Swan", originado a partir de então, apresenta como fundamento-chave a forma neoclássica da função de produção, que assume retornos constantes à escala e retornos decrescentes para cada fator de produção, trabalho e capital. No modelo, a economia possui apenas um setor que é fechado, cujo produto é um bem homogêneo, ou consumido, ou investido, com a taxa de investimento igual a uma taxa de poupança dada exogenamente. O crescimento da população assim como o crescimento da força de trabalho também são exogenamente determinados e, por simplicidade, constantes.

Segundo esse modelo, o processo de acumulação de capital – ou seja, o nível de investimento – assume papel fundamental na determinação do nível de renda do País. O nível de investimento exigido é aquele que mantém a relação capital-trabalho constante. Nesse caso, o investimento em bens de capital precisa suplantar a quantidade necessária para cobrir sua depreciação e a entrada de novos trabalhadores, e esse nível de investimento conduz a sociedade ao crescimento de estado estacionário, *steady state*.

No ponto de *steady state*, o estoque de capital *per capita* fornece o produto que gera poupança e investimento suficientes para que o estoque de capital, o consumo e o produto cresçam à mesma taxa que a população e a oferta de trabalho. Na ausência de progresso técnico, os valores *per capita* são constantes. O crescimento no estado estacionário se refere, portanto, ao crescimento equilibrado de forma que não induza a variações nos preços relativos. Em outras palavras, a variação da razão capital/trabalho no modelo conduz a uma variação na produtividade marginal do capital e do trabalho que não proporciona uma alteração nos preços relativos da economia.

Uma previsão bastante explorada dos modelos derivados da abordagem Solow-Swan é a hipótese de convergência condicional da renda, que provém da suposição de retornos decrescentes para o capital. Segundo tal hipótese, quanto menor o nível inicial do PIB

real *per capita,* relativamente à posição de longo prazo – ou de estado estacionário – maior sua taxa de crescimento. A convergência é condicional porque os níveis de *steady state* do capital por trabalhador e do produto por trabalhador dependem da taxa de poupança, da taxa de crescimento da população e da posição da função de produção, características que variam entre os países.

O processo de acumulação de capital físico assume papel importante à medida que o investimento em máquinas e equipamentos eleva a renda *per capita* e acelera o crescimento dos países. Além disso, políticas que alteram a parcela da renda referente à poupança também auxiliam no processo de aceleração do crescimento e conduzem o sistema à trajetória de crescimento equilibrado. Como as taxas de poupança e de crescimento da população variam entre os países, países diferentes alcançam diferentes estados estacionários. Nessa perspectiva, quanto maior a taxa de poupança, mais rico é o país e, quanto maior a taxa de crescimento da população, mais pobre o país será (Mankiw et al., 1992).

Entretanto, apesar da relevância do investimento em capital físico para alcançar a relação capital por trabalhador do *steady state,* uma vez concluído o período de transição entre os estados estacionários, o modelo prevê que o aumento permanente da taxa de crescimento se sustentará por períodos mais longos unicamente, por meio de mudanças no nível de tecnologia, que, no caso, consiste em uma variável exógena ao modelo. Dada a hipótese de retornos marginais decrescentes para o capital, seria impossível manter uma acumulação de capital físico *per capita* sem a atuação do progresso tecnológico, que seria o responsável por contornar o efeito dos rendimentos decrescentes, mantendo o crescimento do produto *per capita.*

Seguindo a abordagem Solow-Swan, Cass (1965) e Koopmans (1965) desenvolveram um modelo em que a taxa de poupança não é constante, mas sim uma função do estoque de capital *per capita.* Os autores retomaram a análise de Ramsey sobre a otimização do consumo, a qual incorpora ao modelo a taxa de poupança, que passa a ser endógena. Os resultados encontrados pelos autores são similares aos de Solow e Swan, em que as taxas de crescimento das variáveis

por unidade de trabalho são nulas no estado estacionário, sendo o crescimento *per capita* dependente da taxa de progresso tecnológico, a qual permanece exógena ao modelo.

Em resumo, um fator-chave da teoria neoclássica é que o crescimento sustentado do produto *per capita* não ocorre, a menos que haja deslocamentos na função de produção resultantes do progresso técnico exogenamente determinado. Assim, a taxa de progresso técnico determina a taxa de crescimento de longo prazo.

Apesar da relevância, durante muito tempo, houve certa resistência por parte dos autores em inserir a variável tecnologia no modelo. A dificuldade de inclusão de uma teoria da inovação tecnológica na estrutura neoclássica se dá essencialmente porque os pressupostos de concorrência perfeita não podem ser mantidos, uma vez que novas ideias consistem em bens não rivais que adquirem aspectos de bens públicos. Assim, para que fosse possível a inclusão da variável tecnologia, até então exógena ao modelo, seria necessário abandonar o pressuposto de retornos constantes à escala e começar a pensar que os retornos à escala tendem a ser crescentes, se as ideias não rivais são incluídas como fator de produção, o que vai de encontro com o pressuposto de concorrência perfeita.

Desse modo, apesar de tecnicamente bem-sucedidos, os modelos neoclássicos de crescimento econômico perderam fôlego, de forma efetiva, no início dos anos 70, principalmente, por sua clara deficiência na aplicação empírica.

Nos anos 80, a teoria de crescimento econômico voltou a experimentar um novo "*boom*", principalmente, a partir dos trabalhos de Romer (1986) e Lucas (1988). Romer (1986) trabalhou com elementos fornecidos essencialmente por Arrow (1962) e Sheshinski (1967), a fim de introduzir o avanço tecnológico na estrutura competitiva dos modelos neoclássicos (Barro & Sala-I-Martin, 1995). Em seu trabalho, o autor distingue os retornos privados do investimento de seus retornos sociais, sendo que os retornos privados podem ser decrescentes, mas os retornos sociais – que refletem *spillovers* de conhecimento ou outras externalidades – podem ser constantes ou crescentes (Barro, 1990). Por sua vez, o modelo de crescimento de

Lucas (1988) enfatizou os efeitos da qualificação do indivíduo sobre a produtividade, o que compensa o declínio da produtividade marginal do capital. Tais trabalhos reacenderam o interesse pela teoria de crescimento com a incorporação das teorias de P&D e competição imperfeita na estrutura sugerida por Solow-Swan (1956) e Cass-Koopmanss (1965). Nessa linha, uma diferença crucial dos novos modelos em relação aos modelos neoclássicos foi a incorporação do determinante da taxa de crescimento de longo prazo no modelo; o que originou a denominação de "modelos de crescimento endógeno".

Segundo Barro (1990), os modelos recentes de crescimento econômico geram crescimento de longo prazo sem a dependência de variáveis exógenas importantes, como tecnologia e população. Além disso, nos modelos de crescimento endógeno, os retornos do investimento não são necessariamente decrescentes. Conforme Barro & Sala-i-Martin (1995), os *spillovers* de conhecimento e os benefícios externos do capital humano desempenham papel crucial no processo, uma vez que ajudam a evitar a tendência de retornos decrescentes à acumulação de capital.

Em geral, nos modelos de crescimento endógeno, a taxa de progresso tecnológico é afetada por investimentos em P&D, e estes são recompensados por alguma forma de poder de monopólio *ex post*. Entretanto, segundo os referidos autores, as distorções relacionadas à criação de novos métodos de produção conduzem a uma taxa de crescimento que não é ótima no sentido de Pareto, já que os *spillovers* gerados consistem em uma forma de externalidade. Daí a incompatibilidade entre o produto gerado pelos fatores em um contexto de retornos crescentes, que é maior que a contribuição marginal dos mesmos.

Sob esse contexto, tais estruturas teóricas abrem espaço para implicações de políticas públicas, uma vez que a taxa de crescimento de longo prazo dos países depende de atitudes governamentais tais como taxação, "poder de execução" (*enforcement*) das instituições, fornecimento de serviços de infraestrutura, proteção da propriedade intelectual e regulação do comércio internacional e dos mercados financeiros, entre outros aspectos da economia. Em suma, o gover-

no tem grande poder de influência sobre a taxa de crescimento dos países.

Efetivamente, as pesquisas recentes sobre crescimento econômico dão mais ênfase às implicações empíricas do que aos modelos desenvolvidos nos anos 50 e 60. Diversos trabalhos procuraram testar, empiricamente, os resultados obtidos pelos modelos teóricos.

Barro (1990) construiu um modelo que incorpora os gastos do governo financiados por impostos na função de produção da economia. Em um trabalho posterior, Barro (1991) introduziu na discussão de crescimento econômico, um modelo que testa, de forma empírica, a influência de diversos fatores em seu período inicial sobre a taxa de crescimento de uma *cross-section* de países. Nesse último modelo, Barro certificou-se de que a taxa de crescimento do PIB *per capita* real é positivamente relacionada com o capital humano inicial e negativamente relacionada com o nível inicial do PIB *per capita*. Além disso, países com alto nível de capital humano também possuem taxas de fertilidade menores e maior participação de investimento físico no PIB total. O autor testou, ainda, a relação entre crescimento e participação dos gastos com consumo governamental no PIB e verificou que este é inversamente relacionado àquele. Por fim, as taxas de crescimento econômico dos países mostraram-se positivamente relacionadas às medidas de estabilidade política e inversamente relacionadas à *proxy* para distorções no mercado.

É razoável supor uma estreita ligação entre os fatores que determinam o crescimento de um país com aqueles que o fazem em relação ao crescimento de regiões de um mesmo país. Quanto aos últimos, as diferenças na tecnologia, nas instituições e nas preferências são provavelmente menores. Os agentes, firmas e consumidores tendem a ter acesso a tecnologias similares e possuem costumes e preferências parecidos. Além disso, como a legislação geral, os costumes e a língua são os mesmos e não existem barreiras legais à mobilidade dos fatores, esta tende a ser menor entre regiões de um mesmo país.

No entanto, os estudos *a la* Barro (1990,1991) têm sido amplamente criticados sob o ponto de vista econométrico. Autores, como Lee et al. (1997), argumentam que os estimadores são viesados e

que os testes de significância que usam a estatística *t* não são válidos. Além disso, tais trabalhos desconsideram um elemento-chave para a construção de modelos que envolvem estados e/ou municípios, a influência da aglomeração de pessoas e firmas na geração de externalidades geográficas, posto que tal influi, de forma direta, sobre os retornos marginais dos fatores de produção de uma determinada localidade.

A questão das externalidades espaciais e o crescimento das cidades

O estudo da aglomeração de firmas e pessoas em uma determinada localidade vem sendo enfrentado, há um tempo, por autores como Von Thünen (1826), Marshall (1920), Christaller (1933), Lösch (1954) e Jacobs (1969), que buscaram explicar a dinâmica da localização e sua associação com a existência de aglomerações e formação de cidades. A questão central enfrentada por esses autores relaciona-se ao porquê da existência de aglomeração de pessoas e firmas no espaço. A hipótese principal remete aos retornos crescentes à escala, que surgem a partir de economias de aglomeração, isto é, supõe-se que o aumento no número de trabalhadores e firmas, em uma localidade, gera um aumento mais que proporcional no produto dessa região.

O modelo da cidade isolada de Von Thünen introduz a questão ao discutir a dinâmica da localização baseada no uso da terra e nos custos de transporte envolvidos com produção e comercialização. Uma das contribuições mais relevantes de seu modelo é a introdução do conceito de fatores desaglomerativos, em que os custos de congestão exercem um papel de contrapeso das forças aglomerativas. A base do modelo consiste no diferencial entre os custos de transporte de produtos localizados em diferentes pontos do espaço. A presença de produtores mais próximos do centro urbano que, no modelo, é suposto único, favorece o surgimento de uma espécie de monopólio no mercado de terras e produz um "sobrelucro" advindo do baixo

custo do transporte. Por sua vez, o monopólio no mercado de terras influencia, diretamente, a renda fundiária, que varia inversamente com a distância ao centro urbano, formando um gradiente espacial de renda.

O modelo de Von Thünen foi importante também porque abriu caminho para trabalhos, tais como Alonso (1964) e Henderson (1974), os quais formaram a base de sustentação da corrente conhecida como Economia Urbana (*Urban Economics*).

Marshall (1920), da mesma forma, trabalhou com a questão regional e identificou duas fontes para as economias geradas pelo aumento na escala de produção: (1) economias de escala internas às firmas e (2) economias de escala externas às firmas, porém internas ao setor de atividade. Para o autor, existem, essencialmente, três ordens de vantagens em instalar indústrias localizadas, a saber: (1) o mercado de trabalho especializado, (2) o surgimento de indústrias subsidiárias (efeitos de encadeamento) e (3) interatividade de segredos e novas ideias relacionadas à atividade produtiva (*spillovers* de conhecimento). Desse modo, Marshall introduziu o conceito de economias externas e sua relação com as vantagens de se produzir em um distrito industrial. A *tríade marshalliana* das economias externas, como ficou conhecida, mostrou-se notoriamente difícil de ser modelada, mas avançou na questão do porquê as cidades e regiões comerciais centrais existiam.

Em seu trabalho, Henderson (1974) aproveitou as considerações de Von Thünen e Marshall e construiu um modelo que tratava a economia como um sistema urbano, uma coleção de cidades. O autor apontou para a existência de forças *centrípetas* e *centrífugas* que agem, mutuamente, no sentido de escrever o desenvolvimento histórico de uma determinada cidade e/ou região. A tensão existente entre fatores aglomerativos, como economias de escala, e desaglomerativos, como custos de transporte, são a principal justificativa de Henderson para explicar a dinâmica do processo de desenvolvimento dos espaços urbanos.

O trabalho de Jacobs (1969) contrapõe-se, na essência, às ideias de Marshall (1920), uma vez que a autora defende que a especiali-

zação é uma fonte de crescimento limitada e enfatiza para o papel da diversidade das atividades econômicas como fonte do crescimento urbano. Jacobs acredita na inovação como fonte principal de crescimento das cidades. Segundo ela, a inovação surge como novo produto ou serviço que cria novas divisões de trabalho e proporciona novas fontes de criação. Assim, a diversidade das relações de trabalho cria um processo autorreforçador para a geração e fortalecimento do processo de inovação de uma cidade.

De maneira geral, as ideias a respeito das *Economias de Localização* estão associadas ao trabalho de Marshall (1920) e referem-se ao ganho advindo das economias de escala externas às firmas, porém internas à indústria como um todo. Por sua vez, o termo *Economias de Urbanização*, geralmente, associa-se às considerações fornecidas por Jacobs e refere-se às economias externas às firmas, mas internas ao centro urbano.

O modelo da cidade isolada de Von Thünen também serviu de inspiração para uma corrente de teorias da localização, conhecida como Ciência Regional (*Regional Science*). Segundo Fujita et al. (1999), a Ciência Regional tratou de questões que a Economia Urbana desprezou, principalmente, quanto à questão de onde as cidades se formam e a relação espacial entre elas.

Christaller (1933) e Lösch (1940) também desenvolveram um modelo que buscou oferecer uma resposta à questão sobre como as economias de escala e os custos de transporte interagem para produzir uma economia espacial. Na Teoria da Área Central, como ficou conhecido o modelo de Christaller, o autor refere-se ao surgimento de um entrelaçado de áreas principais que surgem com o equilíbrio entre as forças aglomerativas e desaglomerativas. As áreas centrais formam uma hierarquia, com cada grupo de cidades-mercado fazendo parte de um centro administrativo maior. Lösch deu forma a esse sistema de áreas centrais com a afirmação de que, para minimizar os custos de transporte, em determinada densidade de áreas centrais, as áreas de mercado deverão ser hexagonais, e que esse sistema é Pareto eficiente.

Pred (1966) seguiu a tradição da ciência regional e formulou sua teoria por meio da distinção das atividades econômicas de uma região

em dois tipos: primeiro, as atividades que satisfazem a demanda externa e, segundo, as atividades destinadas ao mercado local. A ideia principal do modelo de Pred é de que as atividades voltadas à exportação consistem na base da economia de uma região e que o comportamento das demais atividades é associado ao comportamento das primeiras, crescendo ou se retraindo dependendo do desempenho da base exportadora.

Todavia, apesar de todo o instrumental fornecido pela ciência regional, principalmente, quanto à análise prática, aquela não foi capaz de produzir uma estrutura consistente para os modelos que essa ciência propunha. Tal fato só foi possível com a introdução dos modelos de concorrência imperfeita na estrutura de mercado dos modelos regionais, mais especificamente o modelo Dixit-Stiglitz de concorrência monopolista. Em linhas gerais, o modelo Dixit-Stiglitz preserva os resultados de equilíbrio geral do modelo neoclássico, gerando retornos crescentes a partir das preferências, no caso dos consumidores, ou demandas por variedades, no caso das firmas. Dessa forma, o modelo tornou possível tratar o problema da estrutura de mercado, pois trouxe a questão dos retornos crescentes ao nível da empresa individual e, não somente, tratou-os como fatores puramente externos às empresas.

Essa ligação do modelo Dixit-Stiglitz com a teoria da localização clássica gerou uma perspectiva valiosa sobre como as economias evoluem no espaço. Uma sistematização mais consistente pôde ser construída e permitiu, de alguma forma, a modelagem de uma estrutura de mercado de concorrência imperfeita, associada com o processo por meio do qual uma estrutura espacial organizada surge e se mantém.

Tal perspectiva ganhou dimensão na teoria econômica *mainstream*, principalmente, nos trabalhos de Krugman (1991), Fujita et al. (1999) e Fujita & Thisse (2002), que foram os precursores da nova corrente de pensamento, conhecida como "Nova Geografia Econômica" (NGE). A NGE forneceu meios para lidar com a questão de modelagem sob concorrência imperfeita, a qual, em se tratando de espaço, torna-se elemento-chave, dada a natureza concentradora dos

retornos crescentes à escala. Contudo, a principal contribuição dessa literatura consiste na microfundamentação do comportamento das firmas e dos indivíduos.

Krugman (1991) sugere a primeira versão do modelo centro-periferia e ressalta, por exemplo, o papel da teoria da concorrência imperfeita no tratamento das questões relacionadas à aglomeração de atividades produtivas no espaço. Fujita et al. (2001) tentam explicar questões de localização, tamanho e crescimento das cidades, ao assumirem um comportamento de concorrência perfeita para o setor agrícola, concorrência monopolística para o setor manufatureiro e custos de transporte do tipo *iceberg*.[1]

Uma questão central enfrentada pela NGE refere-se aos incentivos que levam pessoas e firmas a aglomerarem-se em poucos pontos do espaço, mesmo com todas as ineficiências típicas dos grandes centros, como congestionamento, criminalidade e poluição. Nesse sentido, uma das contribuições mais relevantes dessa corrente é a ideia de que a distribuição das atividades depende do resultado de forças contrárias. Sob a visão da NGE, a interação entre externalidades positivas, forças centrípetas, que levam à aglomeração das atividades, e externalidades negativas, que levam a uma dispersão das atividades entre as regiões, resulta em um nível ótimo de concentração econômica.

A perspectiva adotada neste trabalho é de que as externalidades positivas elevam o nível de produtividade de uma determinada região através dos *spillovers* advindos da proximidade de pessoas e firmas. Por sua vez, a elevação da produtividade influencia as taxas de crescimento do emprego e dos próprios centros urbanos.

Conforme os trabalhos de Glaeser et al. (1992) e Glaeser et al. (1995) busca-se, neste estudo, abordar a questão do crescimento de cidades por meio de uma perspectiva dinâmica, na qual, as economias

---

1 Por custos de transporte do tipo *iceberg* entende-se que parte do bem transportado é consumido com o próprio processo de transporte, ou seja, a mercadoria se derrete ao ser transportada, em analogia ao avanço de um *iceberg*. Para maiores detalhes, ver Samuelson (1954).

de aglomeração, sejam elas advindas de economias de localização ou de urbanização, são consideradas tanto em sua extensão geográfica quanto temporal. O caráter geográfico refere-se à atenuação da interatividade dos agentes à medida que estes se tornam mais distantes; já o caráter temporal diz respeito à possibilidade de o comportamento passado dos agentes influenciar o nível atual de produtividade.

As economias estáticas da tradição regional clássica são relevantes para explicar o padrão de localização industrial das cidades – o grau de especialização ou diversificação –, mas não são capazes de elucidar o crescimento de maneira estrita. Marshall e Jacobs fornecem *insights* interessantes às teorias dinâmicas à medida que tratam de economias de localização e urbanização. Tais conceitos baseiam-se em *spillovers* tecnológicos e explicam, essencialmente, o crescimento urbano.

Com o decorrer do tempo, o avanço tecnológico dos meios de comunicação e de transporte alterou a importância relativa da localização geográfica sob o ponto de vista econômico, o que tornou ainda mais complexo o estudo da relação entre proximidade geográfica e dinâmica urbana.

Nessa linha, ao buscar a identificação dos determinantes do crescimento econômico dos municípios paulistas, este estudo adota, como referência, o trabalho de Glaeser et al. (1995), que desenvolveram um modelo para o crescimento populacional e da renda do trabalho em municípios norte-americanos. Acrescentam-se ao modelo, todavia, considerações teóricas da NGE, por meio de ferramentas fornecidas pela econometria espacial, no intuito de quantificar a importância da localização no desempenho de crescimento dos municípios. Assume-se, portanto, que o processo de conexões entre os municípios se autoalimente e resulte na concentração de atividades em determinadas regiões em detrimento de outras.

Na literatura empírica, a análise da influência das externalidades espaciais no crescimento econômico e populacional, em geral, é feita por meio do instrumental fornecido pela econometria espacial, principalmente, a partir do trabalho de Anselin (1988). Os métodos fornecidos pela econometria espacial já foram aplicados em questões de crescimento econômico nas esferas microrregional (Lim, 2003),

regional (Fingleton, 1999), estadual (Rey & Montouri, 1999) e internacional (Moreno & Trehan, 1997). No Brasil, Magalhães et al. (2000), Magalhães (2001) e Silveira Neto (2001), entre outros, estudaram o caso dos estados brasileiros, levando em consideração a existência de *spillovers* espaciais de crescimento. Por sua vez, no que se refere a municípios e microrregiões, Pimentel & Haddad (2004) e Resende (2005) analisaram o caso dos municípios mineiros. Oliveira (2005) estudou o Estado do Ceará, e Monastério & Ávila (2004) utilizaram a econometria espacial para analisar o crescimento econômico de microrregiões do estado do Rio Grande do Sul entre 1939 e 2001.

Conforme preconizado pela primeira lei da Geografia, conhecida como Lei de Tobler,[2] pressupõe-se que microrregiões, bem como municípios, possuam um potencial de influência mútua maior do que as regiões mais abrangentes, como estados e países. Assim, ao estender as análises clássicas de crescimento ao escopo de microrregiões, faz-se necessário um cuidado especial em função de maior interatividade, visto que determinados conjuntos de municípios possuem distâncias relativamente pequenas entre si.

Um modelo de crescimento econômico para os municípios

Adotando pressupostos estilizados na literatura da localização, Glaeser et al. (1995) elaboraram um modelo e testaram-no empiricamente, de forma a relacionar o crescimento de 203 cidades norte-americanas com suas características no período inicial, em 1960. A hipótese basilar que permeia o referido trabalho é que as externalidades positivas geradas pela aglomeração de trabalhadores e firmas em uma determinada cidade elevam a produtividade das economias locais e influenciam, com isso, as taxas de crescimento do emprego e dos próprios centros urbanos.

2 "Everything is related to everything else but nearby things are more related than distant things" (Tobler, 1970, p. 236).

Em conformidade com o trabalho de Glaeser, este trabalho também adota o crescimento populacional das cidades como a principal medida para o crescimento econômico dos municípios. Nesse caso, considera-se que o crescimento populacional funciona como uma *proxy* para a variável crescimento do emprego. Não seria adequado medir o crescimento dessa forma se, por exemplo, fosse um estudo entre países, como é o caso do trabalho de Barro (1991). Isso ocorre, porque, grande parte do crescimento populacional está relacionada às diferentes taxas de natalidade e mortalidade. Além disso, há limitações claras à mobilidade da população no caso de países, o que não acontece, em geral, entre cidades de um mesmo país. Em relação ao crescimento econômico dos municípios, Glaeser et al. afirmam que:

> Entre cidades, o crescimento populacional captura a extensão pela qual estas estão se tornando habitats e mercados de trabalho crescentemente atrativos. O crescimento da renda é uma medida natural do crescimento da produtividade entre os países porque o trabalho é imóvel. Quando o trabalho é móvel, como é o caso das cidades norte-americanas – e também entre estados norte-americanos – a situação é radicalmente diferente. Dentro da economia dos Estados Unidos, a migração responde fortemente ao crescimento das oportunidades. (Blanchard & Katz, 1992) (idem, 1995, p.127)

Desse modo, admite-se o crescimento populacional como uma medida mais apropriada da prosperidade dos municípios, sobretudo, quando se trata de municípios do mesmo estado. Além disso, o crescimento da renda captura declínios na qualidade de vida, o que constitui, portanto, uma medida menos direta do sucesso urbano (idem). Pred (1966) abordou a questão à medida que tratava, em seu modelo, do crescimento demográfico como uma consequência do sucesso urbano e que, posteriormente, funcionava como um fator adicional pela via dos *spillovers* de conhecimento.

No sentido pensado por Glaeser et al., as cidades são consideradas como economias separadas, mas completamente abertas, com livre mobilidade de trabalho, capital e tecnologia. Nesse caso, a tecno-

logia é tratada como um bem público que é livremente acessível e, portanto, não varia entre as regiões. Assim, o crescimento não pode ser explicado por diferenças nas taxas de poupança, participação do capital, taxa de depreciação ou por algum tipo de dotação exógena de mão de obra. Além disso, as cidades são unidades econômicas mais especializadas e menos arbitrárias do que, por exemplo, estados nacionais, sendo que faz mais sentido estudar o movimento de recursos e convergência entre cidades do que entre estados. Conforme tais premissas, as cidades diferem apenas no nível de produtividade e qualidade de vida (idem).

Combinando esses pressupostos, a função de produção utilizada é do tipo Cobb-Douglas, dada por:

$$A_{i,t} f(L_{i,t}) = A_{i,t} L_{i,t}^{\sigma} \quad (2.3.1)$$

Na equação acima, $A_{i,t}$ capta o nível de tecnologia da cidade i no tempo t, enquanto $L_{i,t}$ é a população da mesma cidade no mesmo período; o coeficiente da função de produção σ é suposto constante para todo o país. Assim, como na maioria dos modelos de crescimento, o modelo adotado desconsidera a heterogeneidade da mão de obra, o que pressupõe, dessa forma, trabalho homogêneo.

Tem-se que, no equilíbrio, a renda do trabalhador ($W_{i,t}$) se iguala à produtividade marginal do trabalho:

$$W_{i,t} = \sigma A_{i,t} L_{i,t}^{\sigma-1} \quad (2.3.2)$$

Ao assumir liberdade de migração entre as cidades, asseguram-se utilidades constantes através do espaço em um ponto do tempo, sendo que a utilidade total é dada pelo salário do trabalhador multiplicado por um índice de qualidade de vida. Assume-se que tal índice é uma função monotonicamente inversa ao tamanho dos municípios:

$$\textit{Qualidade de vida} = Q_{i,t} L_{i,t}^{-\delta} \quad (2.3.3)$$

Sendo que $\delta > 0$. Denota-se que o índice de qualidade de vida engloba o efeito de diversos fatores, inclusive crime, preço dos imóveis e congestionamento. Dessa forma, a utilidade total de um potencial imigrante da cidade i é:

$$\text{Utilidade} = \sigma A_{i,t} Q_{i,t} L_{i,t}^{\sigma-\delta-1} \tag{2.3.4}$$

Portanto, a partir da expressão (2.3.4), pode-se inferir que, para cada cidade:

$$\log\left(\frac{U_{i,t+1}}{U_{i,t}}\right) = \log\left(\frac{A_{i,t+1}}{A_{i,t}}\right) + \log\left(\frac{Q_{i,t+1}}{Q_{i,t}}\right) + (\sigma - \delta - 1)$$

$$\log\left(\frac{L_{i,t+1}}{L_{i,t}}\right) \tag{2.3.5}$$

Assumindo que:

$$\log\left(\frac{A_{i,t+1}}{A_{i,t}}\right) = X'_{i,t}\,\beta + \varepsilon_{i,t+1} \tag{2.3.6}$$

$$\log\left(\frac{Q_{i,t+1}}{Q_{i,t}}\right) = X'_{i,t}\,\theta + \xi_{i,t+1} \tag{2.3.7}$$

Nas quais $X_{i,t}$ é um vetor das características das cidades no tempo t que determina tanto o crescimento da qualidade de vida em uma determinada cidade quanto o crescimento de seu nível de produtividade. Combinando (2.3.5), (2.3.6) e (2.3.7) e fazendo algumas manipulações algébricas, tem-se que:

$$\log\left(\frac{L_{i,t+1}}{L_{i,t}}\right) = \frac{1}{1+\delta-\sigma} X'_{i,t}(\beta + \theta) + \chi_{i,t+1} \tag{2.3.8}$$

$$\log\left(\frac{W_{i,t+1}}{W_{i,t}}\right) = \frac{1}{1+\delta-\sigma} X'_{i,t}(\delta\beta + \sigma\theta - \theta) + \omega_{i,t+1} \tag{2.3.9}$$

Sendo que $\chi_{i,t}$ e $\omega_{i',t}$ são termos de erro não correlacionados com as características urbanas.[3] Como resultado, de acordo com os autores, as regressões de crescimento do emprego mostram como as variáveis

---

3 Decorre que, $\chi i,t+1 = [-\log(Ut+1/Ut) + \varepsilon i,t+1 + \xi i,t+1]/(1 + \delta - \sigma)$ e $\omega i,t+1 = [(1- \sigma)\log(Ut+1/Ut) + \delta\, \varepsilon i,t+1 + (\sigma - 1)\, \xi i,t+1]/(1 + \delta - \sigma)$.

ao nível das cidades (os X's) determinam a soma da qualidade de vida e do crescimento da produtividade. Nessa mesma linha, as regressões do crescimento da renda do trabalhador podem ser compreendidas como ilustrativas de uma média ponderada do crescimento da produtividade e ($\sigma - 1$) vezes o crescimento da qualidade de vida.

Em síntese, os resultados obtidos em Glaeser et al. (1995) foram que o crescimento da renda e da população moveram-se conjuntamente, e ambos mostraram-se positivamente relacionados à escolaridade da população no período inicial, negativamente relacionados ao desemprego inicial e negativamente relacionados à participação inicial do emprego industrial. Os gastos do governo, com exceção daqueles utilizados com saneamento, não se mostraram correlacionados ao crescimento, embora se tenha observado que este possuía correlação positiva com o endividamento inicial das cidades.

Em um trabalho semelhante no que diz respeito aos objetivos, mas que compreende um período mais amplo e usa técnicas de Econometria Espacial, Le Gallo & Yrigoyen (2007) examinam o crescimento populacional de 722 municípios espanhóis.

Os autores utilizam uma série de dados bastante ampla e identificam duas fases distintas: de 1900 a 1980, quando se observa divergência entre o crescimento dos municípios, com a concentração da população em grandes cidades, enquanto que o segundo período, que vai de 1980 a 2001, caracteriza-se pela convergência populacional. Tal fenômeno se explicaria por um movimento migratório de "fuga" das grandes cidades, acompanhado de um maior desenvolvimento urbano das cidades pequenas e médias.

Os autores constataram também que a probabilidade de perda de população é cinco vezes maior quando a cidade é cercada por vizinhas que têm população menor, o que confirmaria a hipótese de que as interações espaciais são relevantes para o crescimento das cidades.

No trabalho de Oliveira (2005), é feito um estudo similar para as cidades do Ceará, com base nos censos demográficos de 1991 e 2000. O autor ressalta o papel da educação e urbanização no crescimento das cidades cearenses, assim como a importância da participação do setor público.

No caso específico deste trabalho, o estudo acrescenta às equações do modelo de Glaeser et al. (1995) considerações referentes à influência do espaço nas variáveis adotadas, em linha com os trabalhos de Le Gallo & Yrigoyen (2007) e Oliveira (2005). Tais considerações se fazem necessárias uma vez que o modelo de Glaeser et al. não busca mensurar a presença de custos de transporte de pessoas e insumos que inserem a questão espacial, como fundamental para entender o processo de crescimento e prosperidade dos municípios.

A escolha do modelo espacial mais apropriado é empreendida por meio de técnicas comumente utilizadas no campo da Econometria Espacial, com a utilização de testes específicos, a qual permite a correta especificação do modelo a ser estimado, e torna possível a inclusão de operadores de defasagens espaciais, bem como correções espaciais do termo de erro. O tópico seguinte introduz o tema Econometria Espacial e aborda alguns de seus principais conceitos.

Em síntese, neste capítulo, buscou-se fazer uma breve revisão da literatura de crescimento econômico, além de sua associação com as teorias da localização e com as contribuições da NGE. Nesse sentido, um modelo de crescimento dinâmico foi apresentado. O próximo capítulo traz um resumo das técnicas e métodos de estimação e inferência abordados na econometria espacial clássica, bem como uma discussão mais detalhada sobre a escolha da matriz de pesos espaciais.

2
A abordagem clássica de Econometria Espacial

O termo "Econometria Espacial" foi, inicialmente, introduzido por Jean Paelinck no início dos anos 70 para denominar a área do conhecimento que lida com a estimação e teste de modelos econométricos multirregionais.

A existência de uma área da Econometria denominada de Econometria Espacial se justifica, basicamente, por dois aspectos: o primeiro é a importância da questão espacial inerente à ciência regional, em particular, à economia regional. O segundo é que dados distribuídos no espaço podem apresentar dependência ou heterogeneidade em sua estrutura.

Segundo Lesage (1999), a presença de dependência espacial entre as observações, ou heterogeneidade espacial nas relações modeladas ferem os pressupostos básicos de Gauss-Markov, utilizados, de forma tradicional, em modelos de regressão.

Em termos gerais, heterogeneidade espacial significa que o comportamento econômico não é estável através do espaço, e pode gerar padrões espaciais característicos sob a forma de agrupamentos ao longo do *set* de dados, e variar com a unidade. Dessa forma, os parâmetros variam e podem mudar a forma estrutural do modelo, podendo inclusive, gerar heterocedasticidade com possíveis erros de especificação. Entretanto, Anselin (1988) aponta que, na maioria das vezes, os problemas

gerados pela heterogeneidade espacial podem ser corrigidos com o uso de instrumentos fornecidos pela econometria padrão. Há casos em que o conhecimento teórico da estrutura espacial dos dados pode levar a procedimentos mais eficientes. Além disso, como afirma Resende (2005), o problema torna-se mais complexo naquelas situações em que a heterogeneidade e a autocorrelação estão presentes ao mesmo tempo. Nessas circunstâncias, as ferramentas da econometria padrão são inadequadas e exigem a utilização de técnicas da econometria espacial.

Por sua vez, a dependência ou autocorrelação espacial surge ao se questionar a independência do conjunto de dados coletados. O pressuposto-base para esse tipo de especificação está diretamente associado à primeira Lei da Geografia, na qual todas as informações são relacionadas entre si, porém informações mais próximas estão mais relacionadas do que informações distantes. Assume-se, desse modo, que a proximidade intensifica o processo de conexões entre as unidades espaciais e gera concentração em determinadas localidades em detrimento de outras.

A noção de proximidade, no entanto, é determinada por meio de uma ideia de espaço relativo, ou distância relativa, uma vez que a proximidade não precisa necessariamente estar relacionada à distância entre as localidades. Critérios distintos àquele do sentido euclidiano estrito podem ser considerados, tal como distâncias econômicas, sociais e políticas. O importante é delimitar as regras para uma potencial interação entre as localidades.

No que se refere à metodologia econométrica tradicional, a presença desses efeitos pode tanto requerer alguma modificação na mesma, como pode até invalidá-la. Em alguns casos, faz-se necessária a criação de novas técnicas para o correto tratamento desses efeitos. Como nota Anselin (1988), geralmente, essas questões são ignoradas pela teoria econométrica tradicional e formam o campo específico da Econometria Espacial.

A econometria espacial é importante não apenas quando faz parte da estrutura do modelo, mas também quando ocorrem erros de especificação nas unidades espaciais, os quais podem surgir da não coincidência entre a unidade espacial considerada e a influência

do fenômeno econômico sob consideração, que pode transbordar as fronteiras preestabelecidas. De forma mais específica, a atuação das externalidades pode extrapolar o ambiente de uma cidade, não obedecendo necessariamente seus limites políticos.

Assim como existe a econometria espacial, há também o campo da estatística espacial. Anselin (1988) nota que a distinção entre esses dois campos é sutil, visto que os métodos de uma são amplamente utilizados pela outra. Segundo o autor, mais prático seria deixar a cargo dos próprios pesquisadores referirem seus trabalhos a um ou outro campo.

Em geral, a econometria espacial pauta-se em um modelo ou teoria em particular e tem, como foco, principalmente, a economia regional e urbana, enquanto a estatística espacial trata, de modo primordial, de fenômenos naturais, ligados, principalmente, a campos como a biologia e geologia. A abordagem da econometria espacial consiste basicamente em impor a estrutura do problema por meio da especificação de um modelo *a priori*, ao associá-lo a um teste de especificação com contrapartida em uma hipótese nula. Talvez essa ênfase seja a principal distinção entre a econometria espacial e o campo mais amplo da estatística espacial.

Autocorrelação espacial

Para Anselin & Bera (1998), a autocorrelação espacial pode ser definida como a coincidência entre valores similares e similaridades locacionais. Assim, quando altos ou baixos valores para uma variável aleatória tendem a agrupar-se no espaço, temos o processo de autocorrelação espacial positiva. No entanto, pode acontecer também de as unidades espaciais serem circundadas por unidades com valores significativamente distintos, ou seja, pode ocorrer que altos valores sejam acompanhados por vizinhos com valores baixos, ou vice-versa, processo que se denomina autocorrelação espacial negativa.

Embora os dois processos sejam igualmente importantes e dignos de consideração, a autocorrelação espacial positiva é, sobremaneira, a mais intuitiva, e é encontrada, com maior frequência nos fenômenos

econômicos. Na maior parte das vezes, um processo que apresenta autocorrelação espacial negativa é de difícil interpretação.

Em termos práticos, uma amostra de dados espacialmente autocorrelacionada contém menos informação do que sua contrapartida não autocorrelacionada. Em termos de inferência estatística, essa perda de informação precisa ser levada em conta nos testes de estimação e de diagnóstico. Para Anselin & Bera (1998), esta é a essência do problema de autocorrelação espacial em econometria aplicada.

O problema da autocorrelação espacial tem alguma semelhança com a autocorrelação temporal. De fato, se as regiões de um determinado espaço fossem todas "enfileiradas", de tal modo que só existisse o vizinho da "frente" e o de "trás", (ou, em termos estatísticos, só pudessem apresentar dependência unidirecional) como mostra a figura abaixo, recairíamos em uma situação formalmente idêntica a das séries de tempo e, portanto, todo o tratamento econométrico seria idêntico ao das séries de tempo.

1	2	3	4

Figura 1. Espaço com dependência unidirecional.

Um espaço como o da figura acima é, com evidência, raro de se obter. O caso mais geral é ilustrado pela Figura 2 (embora, não necessariamente, com a mesma regularidade), onde os dados, regiões, estão dispostos em uma superfície bidimensional, e apresentam dependência bidirecional. Assim, a principal diferença entre a dependência temporal e a dependência espacial situa-se, principalmente, na natureza bidimensional e multidimensional da dependência no espaço.

1	2	3
4	5	6

Figura 2. Espaço com dependência multidimensional.

A autocorrelação ou dependência espacial pode ocorrer, basicamente, de duas formas: na variável dependente, ou nos erros. Formalmente, a existência de autocorrelação espacial pode ser expressa pela seguinte condição de momento:

$$Cov(y_i,y_j) = E(y_i,y_j) - E(y_i).E(y_j) \neq 0 \quad \text{para } i \neq j \qquad (3.1.1)$$

Em que y_i e y_j são observações de uma variável aleatória nas localizações *i* e *j* respectivamente. *i* e *j* podem ser pontos, tais como localização de estabelecimentos ou áreas metropolitanas – medidas em latitudes e longitudes – ou unidades de área, tal como países, estados ou municípios (Anselin & Bera, 1998). É evidente que a condição estabelecida por (3.1.1) não é suficiente para que haja um processo de autocorrelação espacial, pois para tal é necessário que a correlação existente entre as observações siga um padrão intuitivo lógico em termos de estrutura espacial.

As consequências da autocorrelação espacial são, em princípio, os mesmos da autocorrelação temporal. Em um modelo de regressão, se os erros são correlacionados entre si (temporal ou espacialmente), os estimadores de mínimos quadrados ordinários são ineficientes, e os estimadores das variâncias serão viesados, o que invalida os testes de significância. Por um lado, para o caso de autocorrelação na variável dependente, as estimativas de MQO são viesadas e inconsistentes, por outro lado, quando a correlação está presente no termo de erro, não há viés, nem inconsistência, mas o estimador de MQO deixa de ser o mais eficiente.

Os processos de autocorrelação espacial guardam analogia com os de séries de tempo, de modo que a situação de autocorrelação serial de ordem 1 pode ser representada da seguinte forma:

$$z_t = {}_t + \rho z_{t-1}, \qquad (3.1.2)$$

em que μ_t é um ruído branco e ρ é o coeficiente de correlação. Em contrapartida, a autocorrelação espacial, também de ordem 1, é mostrada abaixo:

$$z = \mu + \rho \mathbf{W}_1 z \qquad (3.1.3)$$

No caso, z é um vetor *n* por 1 de observações sobre a variável dependente, W_1z é um vetor *n* por *n* de defasagens espaciais para a variável dependente, ρ é o coeficiente autorregressivo espacial, e μ é um vetor *n* por 1 de termos de erro distribuídos aleatoriamente, ou seja, $\mu \sim (0,\sigma^2 I)$. Esse processo é conhecido como SAR (*spatial autoregressive*), onde $\mathbf{W}_1$ é a matriz de conectividade que, em geral, contém relações de contiguidade de 1ª ordem ou funções de distância.[1] Em linhas gerais, $\mathbf{W}_1$ é montada de modo a captar a influência dos vizinhos na variável em consideração. Esse é, portanto, um SAR (1).

Mais genericamente, pode-se ter também um SARMA (*spatial autoregressive moving average*). Segue abaixo um SARMA(1,1).

$$z = \mu + \rho \mathbf{W}_1 z + \theta \mathbf{W}_1 \mu \qquad (3.1.4)$$

Que pode facilmente incluir ordens superiores, e basta, para tal, incluir as respectivas matrizes de conectividade. Por exemplo, o processo abaixo seria um SAR(2).

$$z = \mu + \rho_1 \mathbf{W}_1 z + \rho_2 \mathbf{W}_2 z \qquad (3.1.5)$$

O índice global de Moran (I) é, segundo Anselin & Florax (1995), uma das formas mais amplamente utilizadas de se medir a autocorrelação espacial. Essa estatística varia entre –1 e 1, fornecendo uma medida geral da associação linear (espacial) entre os vetores Z_t no tempo *t* e a média ponderada dos valores da vizinhança, ou *lags* espaciais (WZ_t). Valores próximos de zero indicam inexistência de autocorrelação espacial significativa: quanto mais próximo do valor unitário, mais autocorrelacionado estará. Se o valor dessa estatística for positivo (negativo), a autocorrelação será positiva (negativa). Esse indicador é uma forma de detectar similaridade entre as áreas e é dado por:

$$I = \left(\frac{n}{S_0}\right)\left(\frac{Z'WZ}{Z'Z}\right) \qquad (3.1.6)$$

1 Uma discussão detalhada sobre a matriz de conectividade será realizada no tópico A matriz de pesos espaciais (capítulo 2).

Onde Z é o vetor de n observações para o desvio em relação à média, e S_0 é um escalar igual à soma de todos os elementos de W.

Sendo o valor esperado:

$$E(I) = -\frac{1}{n-1} \tag{3.1.7}$$

Quando a matriz de pesos espaciais é normalizada na linha, ou seja, quando a soma dos elementos de cada linha for igual a um, a expressão poderá ser reescrita, como segue:

$$I = \frac{Z'WZ}{Z'Z} \tag{3.1.8}$$

A estatística I de Moran fornece uma indicação formal do grau de associação linear entre os valores do vetor Z e o vetor espacialmente defasado WZ. Valores maiores do que aqueles esperados, E(I), indicam autocorrelação espacial positiva; negativa, caso contrário.

O diagrama de dispersão de Moran compara os valores normalizados do atributo em uma área com a média normalizada dos vizinhos, o que deriva um gráfico bidimensional de Z(valores normalizados) por WZ (média dos vizinhos). É uma forma de visualizar a dependência espacial e indicar os diferentes padrões espaciais presentes nos dados. O gráfico abaixo representa quatro quadrantes Q_1, Q_2, Q_3 e Q_4 que irão corresponder a quatro padrões de associação local espacial entre as regiões e seus vizinhos.
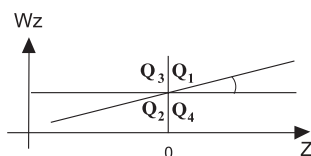
Figura 3. Diagrama de Moran.

O coeficiente I de Moran será a inclinação da curva de regressão de WZ contra Z e indicará o grau de ajustamento. O primeiro qua-

drante, Q_1, conhecido como alto-alto (AA), ou *high-high* – (HH), mostra regiões com altos valores para a variável, valores acima da média, assim como seus vizinhos. O terceiro quadrante, Q_2, geralmente chamado de baixo-baixo (BB) ou *low-low* – (LL), expressa localidades com baixos valores em relação aos atributos analisados, acompanhados por vizinhos que também apresentam baixos valores. O segundo quadrante, Q_3, classificado como baixo-alto (BA) ou *low-high* – (LH), é constituído por baixos valores dos atributos na região estudada, cercada por vizinhos com altos valores. O último quadrante, Q_4, é formado por regiões com altos valores para as variáveis estudadas cercadas por regiões com baixos valores. Este é o quadrante alto-baixo (AB) ou *high-low* (HL).

As regiões de *clusters* com valores similares ocorrem nos quadrantes Q_1 e Q_2 – AA e BB – e apresentam autocorrelação espacial positiva. As regiões identificadas pelos quadrantes Q_3 e Q_4 – BA e AB – apresentam, por sua vez, autocorrelação espacial negativa, ou seja, *clusters* com valores diferentes.

Adicionalmente, a estatística I tem sido usada como um teste para a presença de autocorrelação espacial residual, em linha com a estatística de Durbin-Watson para séries de tempo. Nesse caso, o teste I de Moran é aplicado sobre as estimativas dos erros de uma regressão feita por MQO, com a estatística I observada, comparada com uma distribuição aleatória aproximada por seus momentos, sob a hipótese nula de nenhuma correlação residual. Tiefelsdorf & Boots (1995) fornecem os momentos exatos.

Além da estatística I de Moran, aplicada aos resíduos de uma regressão linear, a presença de algum grau de dependência espacial pode ser verificada por meio de alguns testes específicos, entre eles, o teste de Wald, Razão de Verossimilhança (*Likelihood Ratio – LR*) e através de uma família de testes baseada no Multiplicador de Lagrange (*Lagrange Multiplier – LM*).

Os testes de Multiplicador de Lagrange (LM)[2] são, inclusive, os mais indicados por Anselin (2003) para a escolha da especificação

2 Burridge (1980).

mais adequada. Maiores detalhes sobre testes de especificação e escolha dos modelos serão tratados no tópico Testes de especificação dos modelos espaciais (capítulo 2).

Modelos de regressão com dependência espacial

Segundo Lesage (1999), um modelo autorregressivo espacial mais geral (*spatial autoregressive model* – SAC) pode ser representado da seguinte forma:

$$y = \rho W_1 y + X\beta + \varepsilon,$$

$$\text{Com } \varepsilon = \lambda W_2 \varepsilon + \mu$$

$$\mu \sim N(0,\sigma^2 I_n) \qquad (3.2.1)$$

No modelo acima, y é o vetor nx1 de variáveis dependentes, X é uma matriz nxk de variáveis explicativas, e ε é o termo de erro aleatório normalmente distribuído. W_1 e W_2 são as matrizes nxn de pesos espaciais. Seguindo a definição de contiguidade binária, uma matriz de contiguidade de primeira ordem possui zeros em sua diagonal principal, suas linhas são preenchidas com 0 (zero) nas posições referentes a unidades regionais não contíguas e com 1 (um) naquelas posições vizinhas à unidade que está sendo estudada.[3] ρ, β e λ são parâmetros. É fácil ver que o modelo pode ser reescrito na forma abaixo:

$$(I - \rho W)\,Y = X\beta + (I - \lambda W)^{-1}\,\mu \qquad (3.2.2)$$

O modelo (3.2.1), ou mesmo, sua versão reduzida, em (3.2.2), indica que a dependência espacial se manifesta tanto nas variáveis controladas pelo modelo quanto nas variáveis não controladas. Uma representação esquemática pode ser ilustrada a partir da Figura 4 abaixo.

---

3 No tópico A matriz de pesos espaciais (capítulo 2) são fornecidos alguns exemplos de especificação para a matriz de pesos espaciais.
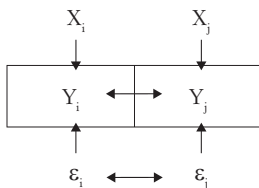
Figura 4. Representação esquemática do modelo SAC.[4]

A Figura 4 ilustra a influência das variáveis explicativas e do termo de erro sobre a variável dependente, sendo que a influência do comportamento dos vizinhos também está presente, tanto na própria variável dependente quanto no termo de erro.

A função logaritmo da verossimilhança (L) para o modelo acima é dada por:

$$\begin{aligned} & \mathrm{L} = C-(n/2)\ln(\sigma^2)+\ln(|A|)+\ln(|B|)-(1/2\sigma^2)(e'B'Be) \\ & e=(Ay-X\beta) \\ & A=(I_n-\rho W_1) \\ & B=(I_n-\lambda W_2) \end{aligned} \qquad (3.2.3)$$

Para a estimação dos parâmetros do modelo SAC, faz-se necessária a otimização do logaritmo da função de verossimilhança. Dessa forma, os estimadores de máxima verossimilhança para ρ e λ requerem que se encontrem os valores dos parâmetros que maximizam o logaritmo da função dada em (3.2.3). Todavia, no sentido de simplificar o problema de maximização, pode-se obter o logaritmo da função concentrada. É possível concentrar a função usando as seguintes expressões para β e σ^2 (Lesage, 1999):

$$\begin{aligned} & \beta=(X'A'AX)^{-1}(X'A'ABy) \\ & e=By-x\beta \\ & \sigma^2=(e'e)/n \end{aligned} \qquad (3.2.4)$$

4 Extraído de Almeida (2007).

Dadas as expressões em (3.2.4), é possível calcular o logaritmo da verossimilhança com os valores de ρ e λ. Os valores dos parâmetros β e σ^2 podem ser calculados como uma função de ρ e λ, e com os dados amostrais de y e X.

Do modelo mais geral, SAC, podem-se derivar modelos distintos ao impor-se restrições sobre os parâmetros. Por exemplo, estabelecendo X = 0 e $W_2 = 0$ tem-se um modelo espacial autorregressivo na forma:

$$y = \rho W_1 y + \varepsilon,$$

$$\varepsilon \sim N(0,\sigma 2 I_n) \qquad (3.2.5)$$

Aqui, o vetor de variáveis y é expresso em termos de desvio da média no intuito de eliminar o termo de intercepto do modelo. O modelo (3.2.5) busca explicar a variação em y como uma combinação linear das unidades vizinhas, sem qualquer outra variável explicativa.

Todavia, dois casos particulares do modelo geral chamam mais a atenção, a saber: quando $W_1 = 0$ ou quando $W_2 = 0$, cada um com problemas econométricos específicos. Nota-se que, se ambas forem iguais a zero, então, o modelo recai no modelo clássico de regressão linear.

No caso de W_2 ser igual a zero, tem-se o modelo com defasagens espaciais SAR (*mixed regressive-spatial autorregressive model*),[5] dado por:

$$y = \rho W_1 y + X\beta + \varepsilon \qquad (3.2.6)$$

O modelo apresenta uma variável explicativa, $W_1 y$, que é o valor médio da variável dependente nos vizinhos. Nesse caso, cada localidade é vizinha de seus vizinhos, tal que o efeito dos vizinhos precisa ser tratado como endógeno. É fácil perceber a similaridade do modelo SAR com o modelo de variáveis dependentes defasadas das séries de tempo. Neste, o período de tempo mais próximo importa, enquanto, naquele, os lugares mais próximos possuem maior relevância. O pa-

5 Anselin (1988) denominou esse modelo como "modelo misto regressivo-autorregressivo espacial" (*mixed regressive-spatial autorregressive model*) porque combina o modelo de regressão-padrão com uma variável espacialmente defasada.

râmetro ρ do modelo (3.2.6) mede a influência média das observações vizinhas sobre as observações do vetor y, o que quer dizer que, para o caso de ρ significativo, uma parcela da variação total de y é explicada pela dependência de cada observação de seus vizinhos.

Nota-se que a presença de um termo para a defasagem espacial do lado direito da equação induz a uma correlação dos erros diferente de zero. Além disso, a defasagem espacial para uma dada observação i não é apenas correlacionada com o termo de erro em i, mas também com os termos de erro em todas as outras localidades. Dado que a simultaneidade incorporada no termo W_1y deve ser explicitamente levada em consideração, a estimativa por MQO será viesada e inconsistente, quando se deve utilizar a função de verossimilhança para estimação. Anselin (1988) fornece um método de Máxima Verossimilhança (MV) para estimar os parâmetros desse modelo.

A figura abaixo ilustra a interação presente no modelo SAR.
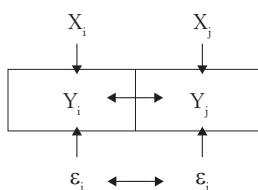
Figura 5. Representação esquemática do modelo SAR.[6]

Como pode ser observado na Figura 5, há uma influência mútua da variável dependente com os seus vizinhos.

Quando uma variável dependente defasada é omitida do modelo de regressão, mas se faz presente no processo gerador dos dados, o problema resultante é similar àquele observado para variáveis omitidas no modelo de regressão linear clássico. Uma alternativa ao método da máxima verossimilhança, nesse caso, seria o uso de variáveis instrumentais, o qual não requer uma suposição de normalidade.

De outra forma, uma maneira de introduzir-se a autocorrelação espacial no modelo de regressão linear é a especificação de uma

6 Extraído de Almeida (2007).

estrutura espacial para o termo de erro. Tal procedimento é necessário quando W_1 é igual a zero. Nesse caso, ocorre um problema de autocorrelação espacial que, do ponto de vista econométrico, tem as mesmas consequências do tradicional problema da autocorrelação temporal: os estimadores de MQO serão ineficientes.

O modelo SEM (*spatial error model*) com autocorrelação espacial no termo de erro é apresentado da seguinte forma:

$$y = X\beta + \varepsilon,$$

$$\text{com } \varepsilon = \lambda W_2 \varepsilon + \mu$$

$$\mu \sim N(0, \sigma 2 I_n) \qquad (3.2.7)$$

y é um vetor nx1 de variáveis dependentes, X representa a usual matriz nxk de variáveis explicativas, e W_2 consiste em uma matriz de pesos espaciais previamente definida. λ e β são parâmetros.

Ao recorrer-se à forma reduzida de (3.2.8), segue-se que:

$$y = X\beta + (I - \lambda W_2)^{-1}\mu \qquad (3.2.8)$$

Neste modelo, a covariância dos erros toma a forma:

$$\in[\varepsilon\varepsilon'] = \sigma^2(I - \lambda W_2)^{-1}(I - \lambda W_2')^{-1} =$$

$$= \sigma^2[(I - \lambda W_2)'(I - \lambda W_2]^{-1} \qquad (3.2.9)$$

Na estrutura da matriz de variância-covariância de (3.2.9), cada localidade é correlacionada com todas as outras localidades do sistema, mas de forma mais intensa com aquelas mais próximas, seguindo a já mencionada Lei de Tobler. O parâmetro de erro espacial, λ, quando significativo, reflete a autocorrelação espacial nos erros ou nas variáveis que foram omitidas do modelo.

Da mesma forma que o processo gerador do modelo de defasagens espaciais, o modelo autorregressivo de erro conduz a uma covariância dos erros diferente de zero para cada par de observações, mas decrescente à medida que aumenta a ordem da contiguidade. Nesse caso, também se deve recorrer à função de verossimilhança.

Segundo Rey & Montouri (1999), quando $\lambda \neq 0$, um choque ocorrido em uma unidade geográfica não se espalha apenas entre seus vizinhos imediatos, mas sim por todas as outras unidades. Uma alternativa é o uso de estimadores do método dos momentos generalizados (GMM), apresentado por Conley (1999).

Segue o quadro esquemático representativo do modelo SEM.
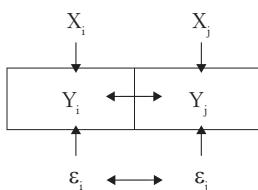
Figura 6. Representação esquemática do modelo SEM.

No caso do modelo SEM, a influência espacial encontra-se nas variáveis omitidas do modelo, como pode ser observado na Figura 6.

Nota-se que, a partir do modelo geral, é possível utilizar a variável W_3X, isto é, a defasagem espacial das variáveis explicativas. Nesse caso, como X é, em princípio, uma matriz de variáveis exógenas, então não há inconveniente sob o ponto de vista econométrico. Esse modelo é conhecido como Modelo Espacial de Durbin (*Spatial Durbin Model* – SDM) (Anselin & Bera, 1998) e assume a forma:

$$y = \rho W_1 y + X\beta - \theta W_3 X\beta + \mu \qquad (3.2.10)$$

$$\mu \sim N(0,\sigma^2 I_n)$$

Sendo que ρW_1 e θW_3 representam as respectivas matrizes de pesos espaciais associadas a seus parâmetros. Nota-se que, em (3.2.10), existe defasagem espacial tanto na variável dependente quanto nas variáveis explicativas. O modelo de *Durbin* pode também ser expresso em termos de variáveis espacialmente filtradas, na forma:

$$(I - \rho W_1)y = (I - \theta W_3)X\beta + \mu \qquad (3.2.11)$$

Este é um modelo de regressão com variáveis dependentes e explicativas espacialmente filtradas e com um termo de erro não autocorrelacionado.

Da mesma forma como foi representado para os outros modelos, segue a ilustração do processo gerado pelo modelo de *Durbin*.
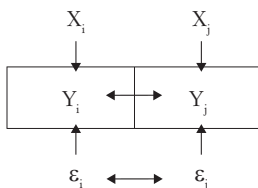
Figura 7. Representação esquemática do modelo SDM.

No último caso, tanto as variáveis explicativas quanto a variável dependente apresentam uma dada estrutura para as unidades espaciais.

Testes de especificação dos modelos espaciais

Como foi visto no tópico anterior, os componentes espaciais do modelo podem aparecer, basicamente, por meio de três formas: (1) na forma de defasagem espacial na variável dependente (Wy), (2) na forma de defasagem nas variáveis explicativas (Wx), ou então (3) como defasagem no termo de erro (Wμ). Tais componentes podem aparecer de forma isolada ou em conjunto. Os testes para modelos espaciais, geralmente, tomam como base a estimação por MV ou por MQO.

Anselin & Bera (1998) enfatizam que, assim como na literatura econométrica clássica, os estágios iniciais da abordagem de econometria espacial foram marcados pela ênfase nas técnicas de estimação. Nesse sentido, Cliff & Ord (1973) desenvolveram a estimação por máxima verossimilhança. Na econometria tradicional, Durbin & Watson (1950, 1951) introduziram a estatística para correlação em modelos de séries de tempo, a qual consistiu no primeiro teste de

especificação aplicado a modelos de regressão. No entanto, outros testes, como: homocedasticidade, normalidade, exogeneidade e forma funcional não tiveram a merecida atenção antes dos anos 80. A estatística descoberta por Rao (1947), que ficou conhecida na literatura como teste do Multiplicador de Lagrange (LM), foi uma exceção e tornou-se amplamente utilizada em função de sua facilidade operacional. Outros testes de natureza "assintótica" também foram desenvolvidos, como o teste de Razão de Verossimilhança (LR) e o teste de Wald.

Apesar de o caminho percorrido pela econometria espacial, em termos de testes de especificação, ter sido muito parecido com o caminho da econometria tradicional, a implementação desses testes se mostrou bastante distinta entre os dois campos de pesquisa. Anselin & Bera (1998) chamam atenção para o fato de que os testes para os modelos de econometria espacial não seguem a forma padrão da maioria dos testes da econometria tradicional, na forma "NR^2" – em que N é o tamanho da amostra, e R^2 é o coeficiente de determinação. Além disso, a possibilidade de defasagem espacial tanto na variável dependente quanto no termo de erro tornam os testes dos modelos espaciais mais complexos.

Conforme já mencionado, a estatística I de Moran surgiu como uma analogia bidimensional ao teste de Durbin-Watson para séries de tempo e, desde então, é a técnica mais utilizada para diagnosticar autocorrelação espacial em modelos de regressão.

A estatística de Moran possui como hipótese nula a inexistência de qualquer forma de dependência espacial, mas não apresenta uma correspondência direta com uma hipótese alternativa particular. Assim, apesar de ser um bom identificador de correlação espacial, o teste não é capaz de distinguir qual estrutura de dependência espacial está presente no modelo.

Recentemente, uma variedade de testes alternativos à estatística I tem sido desenvolvida.[7] Assim como o teste de Moran, outros testes também são baseados nos resultados de uma regressão de

7 Para maiores detalhes ver Anselin & Florax (1995).

MQO clássica, ao apresentarem como hipótese nula a ausência de autocorrelação espacial.

Um modelo mais geral de dependência espacial é o modelo SARMA, demonstrado em (3.1.4). Aqui, acrescenta-se ao modelo uma componente com variáveis explicativas exógenas, conforme Anselin & Florax (1995).

$$y = \rho W_1 y + X\beta + \theta_1 W_2 \mu \qquad (3.3.1)$$

Sendo que as notações permanecem as mesmas das equações anteriores. O primeiro termo $\rho W_1 y$ representa a variável dependente espacialmente defasada, com um parâmetro espacial autorregressivo ρ. O segundo termo do lado direito da equação, $X\beta$, representa a matriz de variáveis explicativas exógenas mais o vetor de parâmetros β. O último termo $\theta_1 W_1 \mu$ refere-se à defasagem no termo de erro, mais o parâmetro θ_1.

Do modelo geral, segue-se que os testes baseados nas estimativas de MQO são aplicados somente a um tipo de dependência, sendo assumida, de forma condicional, a ausência do outro tipo. Assim, a hipótese nula para testar a presença de um processo autorregressivo espacial é H_0: $\rho = 0$, condicionado a $\theta_1 = 0$. Anselin & Florax (1995) chamam atenção para quando essas condições não são satisfeitas, ou seja, quando a presença de uma outra forma de dependência espacial está presente no modelo. Nesse caso, os testes não podem mais ser baseados nos resultados da regressão de MQO, e devem ser levados a cabo por meio das estimativas de MV do modelo espacial apropriado; ou ainda, podem-se utilizar testes robustos que considerem a presença da outra forma de dependência espacial.

No caso deste trabalho, além do I de Moran, dois testes familiares de dependência espacial em modelos de regressão linear são investigados, LM-ERR e LM-LAG. Assim como a estatística de Moran, a família de testes LM utiliza apenas os resultados das estimativas por MQO, sob a luz de uma H_0 de nenhuma dependência espacial. A estatística LM-ERR foi sugerida por Burridge (1980) e é, basicamente, um coeficiente de Moran em escala quadrática. A estatística para o teste apresenta-se da seguinte forma:

$$\text{LM-ERR} = \frac{\left(e'W_1e/s^2\right)^2}{T_1} \tag{3.3.2}$$

Em que $s^2 = e'e/n$ e $T_1 = \text{tr}\left(W'_1 W_1 + W_1^2\right)$, com tr como um operador traço da matriz. A estatística LM-ERR segue uma distribuição χ^2 com 1 (um) grau de liberdade e possui, como hipótese alternativa, a presença de dependência espacial no termo de erro.

O teste LM para a presença de dependência espacial na variável dependente é dado por:

$$\text{LM-LAG} = \left(\frac{e'W_1y}{s^2}\right)^2 \frac{1}{\left(nJ_{\rho-\beta}\right)} \tag{3.3.3}$$

Com $J_{\rho\text{-}\beta} = [T_1 + (W_1X\beta)'M(W_1X\beta)/s^2]$ e $M = I - X(X'X)^{-1}X'$ que é a matriz de projeção usual. A estatística LM-LAG também segue uma distribuição χ^2 com 1 grau de liberdade.

Bera & Yoon (1993) fornecem as versões robustas dos testes LM-ERR e LM-LAG, as quais consideram o efeito da dependência espacial que não é captado pelo teste. O teste LM-EL é o teste LM para dependência espacial no termo de erro, robusto à dependência espacial na variável dependente. O teste é computado da seguinte forma:

$$\text{LM-EL} = \frac{[e'W_1e/s^2 - T_1(nJ_{\rho-\beta})^{-1}(e'W_1y/s^2)]^2}{[T_1 - T_1^2(nJ_{\rho-\beta})^{-1}]} \tag{3.3.4}$$

E a notação permanece a mesma das anteriores, e a distribuição de LM-EL permanece uma χ^2 com 1 grau de liberdade.

O teste robusto para LM-LAG consiste em um teste para defasagem espacial que considera a influência da dependência espacial no erro. O teste LM-LE é definido formalmente, como segue:

$$\text{LM-LE} = \frac{\left(e'W_1y/s^2 - e'W_1e/s^2\right)^2}{nJ_{\rho-\beta} - T_1} \sim \chi^2_{(1)} \tag{3.3.5}$$

Florax et al. (2003) afirmam que os testes robustos do multiplicador de lagrange possuem um poder maior em apontar a alternativa

correta para a especificação do modelo, ao invés de serem adotados os testes LM tradicionais. Adicionalmente, os autores fornecem uma estratégia de especificação híbrida que combina a taxonomia clássica de especificação com o emprego dos testes robustos, como segue:

1. estima-se o modelo inicial $y = X\beta + \varepsilon$ através de MQO;
2. testa-se a hipótese de nenhuma dependência espacial em função de uma defasagem espacial omitida, ou em função de erros espacialmente autorregressivos, utilizando LM-LAG e LM-ERR, respectivamente;
3. se ambos os testes forem não significativos, as estimativas iniciais do passo (1) devem ser usadas como a especificação final; caso contrário, procede-se como sugerido em (4);
4. se ambos os testes são significativos, estima-se a especificação apontada por aquele mais significativo dos dois testes robustos. Por exemplo, se LM-LE > LM-EL, então, estima-se (2) usando LM-LAG. Se LM-EL > LM-LE, então, estima-se (2) usando LM-ERR. De outra forma, procede-se como em (5);
5. se LM-LAG é significante, mas LM-ERR não o é, estima-se (2) utilizando LM-LAG. Caso contrário, procede-se como sugerido em (6);
6. estima-se (2) usando LM-ERR.

Lesage (1999) apresenta um outro teste com base no multiplicador de lagrange que possibilita analisar se a presença do termo de defasagem espacial elimina a dependência espacial presente nos resíduos do modelo de MQO. Essa estatística testa a presença da dependência espacial nos resíduos, condicionada à existência de um parâmetro ρ para a defasagem espacial diferente de zero.

O teste é baseado no seguinte modelo (Lesage, 1999):

$$y = \rho C y + X\beta + \mu$$

$$\mu = \lambda W \mu + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n) \quad (3.3.6)$$

Sendo que o foco do teste é sobre o parâmetro λ.

A estatística para o teste apresenta a seguinte forma:

$$\left(e'We/\sigma^2\right)\left[T_{22}-\left(T_{21}\right)^2\operatorname{var}\left(\rho\right)\right]^{-1}\sim\chi^2_{(1)} \tag{3.3.7}$$

Com $T_{22} = \operatorname{tr}\left(W.*W+W'W\right)$

$T_{21} = \operatorname{tr}\left(W.*CA^{-1}+W'CA^{-1}\right)$

Tem-se que W é a matriz de pesos espaciais escolhida, $A=(I_n-\rho C)$, var(ρ) é a estimativa de MV para a variância do parâmetro ρ no modelo e . * simboliza a operação de multiplicação da matriz elemento por elemento.

Por sua vez, o teste assintótico de Wald não é baseado nos resultados de uma regressão de MQO, mas sim no cômputo das estimativas de máxima verossimilhança do modelo espacial apropriado. A estatística de Wald pode ser utilizada tanto para averiguar a presença de dependência espacial na variável dependente quanto no termo de erro. Contudo, é mais comum encontrar o teste aplicado ao modelo de erro espacial, com H_0: λ=0, sendo a hipótese alternativa o modelo SEM. O teste aplicado ao modelo de erro espacial é definido como:

$$W=\lambda^2\left[t_2+t_3-(1/n)\left(t_1^2\right)\right]\sim\chi^2_{(1)}$$

$$t_1 = tr(W.*B^{-1})$$

$$t_2 = tr(WB^{-1})^2$$

$$t_3 = tr(WB^{-1})'(WB^{-1}) \tag{3.3.8}$$

Em que $B=(I_n-\lambda W)$, com λ sendo a estimativa de MV.

Por fim, o teste de Razão de Verossimilhança (LR) é baseado na diferença entre o logaritmo (log) da verossimilhança do modelo SEM e o log da verossimilhança do modelo de MQO. Dessa forma, Anselin (1988) define o teste como:

$$LR=2[L(\theta)-L(\theta_R)] \tag{3.3.9}$$

Sendo que $L(\theta)$ corresponde ao log da verossimilhança do modelo não restrito – modelo SEM – e $L(\theta_R)$ corresponde ao log da verossimilhança do modelo restrito, ou seja, o modelo de MQO. O teste *LR* é distribuído assintoticamente como uma χ^2 com *q* graus de liberdade.

A matriz de pesos espaciais

Considerando um modelo de regressão linear familiar da forma:

$$y = X\beta + \varepsilon \qquad (3.4.1)$$

A matriz de variância-covariância dos erros, cov[ε ε'], expressa uma covariância espacial quando os elementos fora da diagonal principal são diferentes de zero e seguem uma dada estrutura ou ordenamento espacial (Anselin 2003), que especifica os pares de localidades i-j (com $i \neq j$) cuja covariância será diferente de zero, ou $\in[\varepsilon_i \varepsilon_j] \neq 0$.

Há duas maneiras de encontrar o padrão espacial dessa estrutura. A primeira maneira consiste em especificar, diretamente, a covariância como uma função da distância que separa quaisquer dois pares de localidades. Essa abordagem é comumente empregada em geoestatística, onde as superfícies espaciais são contínuas. Segundo Anselin (2003) tal abordagem requer uma função decrescente para a distância e um parâmetro espacial que assegurem uma matriz de variância-covariância definida positiva.

A segunda forma de encontrar o ordenamento espacial – mais adequada para pontos de observação discretos no espaço – requer a especificação de um processo estocástico que relacione o valor de uma variável aleatória em uma localidade aos valores dessa variável em localidades vizinhas. Assim, em vez de ligar todos os pares por meio de uma função de decaimento da distância, os vizinhos de cada localidade são especificados por meio da chamada matriz de pesos espaciais, W (Anselin, 2003). Dessa forma, para cada ponto do espaço, é definido um conjunto de vizinhança relevante que, potencialmente, interage com ele.

A segunda abordagem aproxima-se mais da realidade dos dados econômicos, uma vez que tal perspectiva é uma extensão do caso tradicional para as séries de tempo, no entanto, para um ordenamento em um espaço bidimensional. De fato, a Econometria Espacial propriamente dita está relacionada aos dados em treliças, pontos discretos no espaço. Uma "treliça" de locações vem da ideia de pontos

espaçados (regiões) ligados a seus "vizinhos", como o exemplo da Figura abaixo.

1	2	3
4	5	6

Figura 8. Dados em treliças.

O ordenamento das informações ao longo do espaço pode ser feito de diversas maneiras. Uma delas é o critério de contiguidade, que reflete a posição de uma unidade em relação às demais unidades no espaço. Medidas de contiguidade necessitam de informações a respeito do tamanho e forma das unidades regionais. Quanto à dependência espacial, pressupõe-se que regiões vizinhas, contíguas, apresentem um grau maior de dependência do que as demais.

Por exemplo, seguindo o critério "Rainha" (*Queen*) de contiguidade,[8] na Figura 8, a região 1 é vizinha das regiões 2, 4 e 5, enquanto a região 5 é vizinha de todas as demais. O critério rainha estabelece que essas relações de vizinhança podem ser representadas pela matriz de conectividade **W** abaixo:

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Obviamente, a matriz W é simétrica, e evidencia que, assim como a região 1 é vizinha da região 2, a região 2 é vizinha da região

8 O critério Rainha considera como vizinhas as unidades que possuem fronteiras ou vértices comuns, em que a unidade vizinha é definida da forma $w_{ij} = 1$, enquanto o elemento, que não possui relação de vizinhança, é definido $w_{ij} = 0$.

1. Além disso, por convenção, a matriz sempre tem zeros em sua diagonal principal.

Pode-se também construir a matriz **w**, que seria a matriz **W** normalizada pelas linhas, isto é, alterada de tal modo que a soma em cada linha seja exatamente igual a 1. Isso é feito simplesmente dividindo o valor de cada elemento da matriz pelo total das linhas.[9] Dessa forma, a soma das influências dos vizinhos é igual para cada unidade em consideração, o que torna possível a comparabilidade. Além disso, com a normalização da matriz **W**, a amplitude de possibilidades dos parâmetros é restrita ao intervalo de -1 a 1.

$$\mathbf{w} = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{pmatrix}$$

A motivação para a normalização da matriz W foi ilustrada por Lesage (1999), da seguinte forma:[10] em primeiro lugar, considera-se a matriz de multiplicação **w** e um vetor de observações de alguma variável associada com as seis regiões a que se chama de y.

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \end{pmatrix} = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/5 & 0 & 1/5 & 1/5 & 1/5 & 1/5 \\ 0 & 1/3 & 0 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 0 & 1/5 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix}$$

A matriz produto $y^* = \mathbf{w}y$ representa uma nova variável igual à média das observações das regiões contíguas.

---

9 Cf. Anselin (1988).

10 Para o caso deste trabalho, adaptou-se a representação esquemática de Lesage ao mapa fornecido pela Figura 7.

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \end{pmatrix} = \begin{pmatrix} 1/3y_2 + 1/3y_4 + 1/3y_5 \\ 1/5y_1 + 1/5y_3 + 1/5y_4 + 1/5y_5 + 1/5y_6 \\ 1/3y_2 + 1/3y_5 + 1/3y_6 \\ 1/3y_1 + 1/3y_2 + 1/3y_5 \\ 1/5y_1 + 1/5y_2 + 1/5y_3 + 1/5y_4 + 1/5y_6 \\ 1/3y_2 + 1/3y_3 + 1/3y_5 \end{pmatrix}$$

Uma das maneiras de quantificar a relação $y_i = f(y_j)$, $j \neq i$, é por meio da matriz de conectividade binária. Considera-se, em ambos os casos, vizinhança de ordem 1. A região 1 não é vizinha (de ordem 1) da região 3, mas elas são vizinhas de segunda ordem, pois a região 1 é vizinha da região 2 que, por sua vez, é vizinha da região 3. Para relações de ordem 2 ou superiores, são necessárias, portanto, diferentes matrizes de conectividade. A matriz **W** de ordem 0 ($\mathbf{W_0}$) é a própria matriz identidade.

Existem outras formas de montar a matriz W. Basta que a forma escolhida considere algum tipo de medida que estabeleça a participação dos vizinhos. Nesse sentido, busca-se aqui testar uma variedade de matrizes de pesos a fim de se identificar aquela que mais se aproxima da verdadeira correlação espacial apresentada pelos dados. Para tal, algumas matrizes W serão testadas, entre elas, a matriz de contiguidade binária.

Como já foi mencionado, a matriz W tem o intuito de captar a estrutura de correlação espacial apresentada pelos dados. Assume-se, dessa forma, uma estrutura específica para o erro, sendo que a literatura de econometria espacial admite uma certa arbitrariedade na seleção da matriz de pesos e permite escolhas *ad hoc* por parte do pesquisador. Conforme Anselin (1988), a escolha apropriada da matriz de pesos espaciais é uma das questões metodológicas atuais mais controversas em econometria espacial.

Dado que as hipóteses sobre a matriz W são feitas *a priori*, a eliminação dos resíduos espacialmente autocorrelacionados não é uma condição suficiente para a eliminação do viés de estimação e pode diferir da verdadeira função. Daí a importância da escolha adequada da matriz

de pesos. Florax & Rey (1995) exploram e enfatizam as implicações da má especificação da matriz W.

Lesage (1999) sugere que o princípio direcionador da escolha da matriz mais adequada deve ser a natureza do problema a ser modelado, sendo relevante, também, o emprego de informações adicionais não fornecidas pela amostra. Abreu et al. (2005) atentam para a importância de fundamentar-se a escolha da matriz de pesos, segundo um conjunto de hipóteses teóricas feitas *a priori*.

As matrizes de pesos espaciais mais tradicionais são construídas a partir de atributos físicos e geográficos, como vizinhança, distâncias geográficas e tempo de deslocamento.

De acordo com a distância geográfica, a construção da matriz W está baseada no ordenamento de um espaço cartesiano representado por latitudes e longitudes. Esse tipo de ordenamento permite calcular as distâncias de quaisquer pontos no espaço. Com relação à dependência espacial, pressupõe-se que o grau de dependência é negativamente relacionado com a distância. Em outras palavras, assume-se que a intensidade da dependência espacial declina à medida que a distância entre as unidades aumenta.

Uma matriz de pesos baseada em distâncias geográficas pode ser calculada por meio do inverso da distância euclidiana, na forma:

$$w^*_{ij} = \frac{1}{\sqrt{(x_i - x_j)^2 + \left(y_i - y_j\right)^2}}, \text{ se } i = j$$

$$w^*_{ij} = 0, \text{ se } i = j \qquad (3.4.2)$$

x_i, x_j, y_i e y_j são as coordenadas dos centroides das unidades i e j.

Nessa especificação, a construção da matriz de pesos é feita com base em um grande círculo entre as regiões, centroides. Entretanto, o emprego do inverso da distância euclidiana ao quadrado é mais usual, uma vez que maiores distâncias são penalizadas mais rapidamente, ou seja, atribui-se maior peso aos vizinhos mais próximos. Tem-se, portanto:

$$w^*_{ij} = \frac{1}{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, \text{ se } i \neq j$$

$$w^*_{ij} = 0, \text{ se } i = j \qquad (3.4.3)$$

A notação permanece a mesma dos conjuntos de equações anteriores.

Uma especificação adicional consiste em construir a matriz de pesos espaciais por meio de uma distância limite, mas com um número fixo, *k*, de vizinhos mais próximos. A matriz W(k) usual é definida, como se segue:

$$w^*_{ij}(k) = 0 \text{ se } i = j, \forall k$$

$$w^*_{ij}(k) = 1 \text{ se } d_{ij} \leq d_i(k)$$

$$w^*_{ij}(k) = 0 \text{ se } d_{ij} > d_i(k) \qquad (3.4.4)$$

Em que $d_i(k)$ é a distância do vizinho de ordem *k*. Segundo Ertur & Gallo (2003), existem diversas vantagens para a preferência dessa matriz em contrapartida à matriz de contiguidade simples. Em primeiro lugar, ela tem vantagem quando ilhas importantes fazem parte da amostra de dados. Um exemplo é o caso da Grã-Bretanha, que seria relegada, no caso de uma análise espacial dos países europeus, porque não possui vizinhos. Em segundo lugar, de acordo com os referidos autores, ao escolher um número fixo de vizinhos, evita-se uma série de problemas metodológicos que surgem quando se permite a variação nesse número.

Tysler (2006) chama atenção para uma matriz de pesos que utiliza um número fixo de vizinhos e que não é construída de forma binária, mas sim através da distância entre os centroides. Essa matriz W é construída da seguinte forma:

$$w^*_{ij}(k) = 0, \text{ se } i = j$$

$$w^*_{ij}(k) = \frac{1}{d_{ij}^{\,2}}, \text{ se } d_{ij} \leq d_i(k)$$

$$w^*_{ij}(k) = 0, \text{ se } d_{ij} > d_i(k) \qquad (3.4.5)$$

Nesse caso, $d_i(k)$ é a distância de corte, isto é, a distância do vizinho de ordem k.

No caso deste trabalho, maior atenção será dada a algumas matrizes em específico, a saber, a matriz de pesos binária, que leva em consideração as fronteiras e vértices comuns, e a matriz formada pelo inverso da distância geográfica com um número fixo de vizinhos. A última foi escolhida, conforme a recomendação de Ertur & Gallo (2003), que, como foi visto, apontam inúmeras vantagens em se utilizar uma matriz que fixa o número de vizinhos.

Case & Rosen (1993) e Conley & Ligon (2002), entre outros, sugerem o uso de pesos baseados na "distância econômica" entre as regiões. Especificamente, Case & Rosen (1993) sugerem usar pesos (antes da padronização) na forma $w_{ij} = \frac{1}{|x_i - x_j|}$, em que x_i e x_j são observações socioeconômicas características da unidade, tais como renda *per capita* ou percentual da população em determinado grupo ético ou racial.

Esses autores utilizam o conceito de similaridade para pressupor uma conexão maior entre as unidades espaciais, ao invés de unidades próximas. Nesse sentido, um município que lidera, economicamente, uma determinada região pode sofrer mais influência de um município líder da região vizinha do que de municípios mais próximos, mas que não possuem economia similar sua.

Na prática, busca-se captar as diferenças existentes entre os municípios para um mesmo indicador socioeconômico, no qual a distância euclidiana invertida é a mais comumente usada.

Conley & Ligon (2002) utilizam uma matriz de distância econômica que discrimina os custos de transporte do capital físico daqueles observados para o capital humano. Para tal, os autores fazem uso dos custos de transporte de encomendas (*United Parcel Service* – UPS) entre capitais de países selecionados, no sentido de medir os custos de transporte do capital físico e os preços de passagens aéreas, buscando capturar os custos de transportar-se capital humano. Os autores procuram, dessa forma, montar a matriz W. Ao levar em conta tal especificação, a razão da distância econômica entre Brasil e Áustria,

por exemplo, e a distância entre Brasil e Austrália é menor do que a mesma razão quando se considera a distância geográfica.

Contudo, os autores admitem a correlação existente entre os custos de transporte propriamente ditos e a distância geográfica, uma vez que, para eles, se o principal impedimento aos *spillovers* consiste no custo de transportar fatores, então parece evidente que a distância geográfica será correlacionada com esses custos. (Conley & Ligon, 2002, p.168)

3
Descrição dos dados

As unidades espaciais adotadas consistem nos municípios paulistas, e as fontes provêm das edições do Instituto Brasileiro de Geografia e Estatística (IBGE) para os anos de 1980 e 2000.[1] Em linha com Glaeser et al. (1995), utilizou-se a taxa de crescimento populacional dos municípios como uma *proxy* para o crescimento econômico. Entretanto, a utilização de dados demográficos gerou problemas de ordem comparativa, uma vez que a quantidade de municípios existentes no Estado de São Paulo não foi constante ao longo dos anos.[2] Dessa forma, constatou-se que, em 1980, havia 571 municípios no Estado contra 645, no ano de 2000. Desse modo, a criação de novos municípios provoca uma distorção na análise, já que a perda de população de um número não negligenciável de municípios poderia decorrer, essencialmente, da criação de novas unidades administrativas. Essa distorção foi corrigida de forma que apenas os

---

1 Extraídos da Base de Dados do Instituto de Pesquisa Econômica Aplicada (Ipeadata).

2 No que se refere ao período em estudo, observa-se uma complicação adicional decorrente da promulgação da Constituição Federal de 1988, que modificou a regulamentação a respeito da criação de novos municípios no País (decreto LC nº 9, de 9 de novembro de 1967), afrouxou os critérios vigentes até então, o que provocou uma verdadeira explosão municipalista.

571 municípios existentes no ano de 1980 foram utilizados. A correção foi feita por meio do agrupamento dos territórios emancipados aos municípios de origem.

A Tabela 1 ilustra a variação populacional dos municípios entre os anos de 1980 e 2000, além de algumas de suas características no ano de 1980. A variação populacional foi representada pela taxa de crescimento média anual da população de cada município. A partir da Tabela 1, é possível ter uma ideia do grau de disparidades na taxa de crescimento entre os municípios paulistas. O Estado apresentou uma variação média de aproximadamente 1,59% a.a., porém com municípios com taxas médias negativas de -2,48% ao longo do período, e municípios que cresceram a taxas anuais médias de 10,53%. O tamanho médio da população das cidades paulistas foi de 43.857 habitantes em 1980, e crescem para uma média de 57.415 em 2000, o que corresponde a um crescimento médio de aproximadamente 31% ao longo do período.

Tabela 1 – Variação da população 1980-2000 e variáveis municipais em 1980: média e desvio-padrão

Variáveis – Abreviações	Média	Desvio-Padrão	Mínimo	Máximo
Variação da população 1980-2000 (em %)	1,59	1,68	–2,48	10,53
Escolaridade da população com mais de 25 anos (em anos de estudo)	2,90	0,76	1,20	5,90
Taxa de analfabetismo da população com mais de 15 anos de idade (em %)	21,96	6,09	4,20	47,10
% de casas com água encanada	75,60	18,09	18,90	100,00
% de casas com iluminação elétrica	82,15	0,16	16,65	100,00
Renda *per capita* (em salários mínimos)	1,37	0,44	0,58	3,46
Índice de Theil	0,43	0,12	0,17	1,15
Esperança de vida ao nascer	59,4	2,55	53,22	66,29
Mortalidade infantil (por mil nascidos vivos)	58,19	12,77	27,73	92,93
Taxa de homicídios (a cada 100 mil habitantes)	6,33	9,52	0	49,66
% do PIB devido à indústria	31,50	21,15	0,72	84,96
% do emprego no setor urbano	63,28	21,77	11,18	100,00

Fonte: IPEA.

No que se refere ao conjunto das variáveis explicativas, utilizou-se, neste trabalho, das variáveis Escolaridade Média e Taxa de Analfabetismo no sentido de obter evidências a respeito do papel da educação no nível do crescimento municipal.[3] Seguindo os trabalhos de Mankiw et al. (1995), Romer (1991), bem como a literatura da NGE, o nível de educação é utilizado como *proxy* para o capital humano, sendo que o pressuposto inicial é que a escolaridade média apresenta uma relação direta com o crescimento, enquanto a taxa de analfabetismo, uma relação inversa. O papel da infraestrutura fornecida pela rede pública[4] é representado pelo percentual de domicílios com iluminação elétrica. A variável renda *per capita* é tradicionalmente utilizada nos modelos de crescimento, assim como nos trabalhos da NGE. Exemplos são os trabalhos de Barro & Sala-i-Martin (1995) e Rey & Montouri (1999). Aqui, ela também foi inserida no modelo, sendo medida em unidades de salário mínimo.

A qualidade de vida oferecida pelos municípios também é captada por meio das variáveis Expectativa de Vida e Mortalidade Infantil (por mil nascidos vivos), que indicam o estado de saúde da população municipal, e a Taxa de Homicídios (a cada 100 mil habitantes). A inspiração para a utilização de tais variáveis vem da NGE.

Um conjunto adicional de termos do modelo segue, de certo modo, o trabalho de Barro (1991) e refere-se aos efeitos que a aglomeração urbana pode trazer ao crescimento dos municípios. Por um lado, tem-se que quanto maior a população do município, tudo o mais mantido constante, maior a probabilidade de surgimento de economias de aglomeração. Nesse caso, a importância do número de habitantes do município é levada em consideração. Busca-se, assim, por meio da variável Logaritmo da População, captar os efeitos positivos advindos da aglomeração de pessoas. No entanto, em linha com Silva Júnior (2007), uma população grande em uma ampla área geográfica pode não trazer os resultados previstos. Por isso, a variável Área Municipal é utilizada como controle.

3 A descrição detalhada das variáveis se encontra no Apêndice 1.

4 Conforme Barro (1990), entre outros.

Por outro lado, uma vez que a concentração populacional elevada pode trazer resultados indesejados, as chamadas deseconomias de escala – ou efeitos de congestionamento, utilizou-se da variável Logaritmo da População ao Quadrado, a fim de se identificarem os efeitos negativos da concentração populacional. Essa abordagem segue, em algum grau, aquela adotada por Ciccone & Hall (1996) e Ciccone (2002).

Em resumo, espera-se valor positivo e significante para o termo linear da população no período inicial, enquanto o termo quadrático deve ter sinal negativo e significante. Em outras palavras, presume-se que a relação entre o crescimento econômico municipal no período e o tamanho da população em 1980 tenha o formato de "U" invertido.

Por fim, a composição da economia municipal é considerada ao se discriminar o papel dos setores da atividade econômica. Para tanto, utilizou-se a variável parcela do PIB relativa ao setor industrial. O percentual da força de trabalho empregada na zona urbana é utilizado para captar possíveis efeitos de economias de urbanização.

Buscando tornar o modelo robusto às influências espaciais, utilizou-se da variável Distância dos Municípios à Capital Estadual. Dessa forma, essa variável funciona como controle e torna os resultados do modelo, a princípio, não sujeitos a esse tipo de influência. Nessa linha, seria possível o uso da variável Custos de Transporte à Capital Estadual. Entretanto, dada a alta correlação existente entre as duas variáveis, a distância à capital foi escolhida de forma arbitrária.

A matriz de correlação, na Tabela 2, auxilia não só na escolha das variáveis, como também na visualização das principais correlações existentes entre as variáveis adotadas. Destaca-se a forte correlação existente entre as variáveis Escolaridade e Renda *per capita*, no valor de 0,82, e Escolaridade e Percentual do Emprego Urbano, com o valor de 0,75. Um outro indicador da interação entre o nível de renda e o nível de educação do município foi o valor de -0,72 de correlação bruta entre as variáveis "Renda *per capita*" e "Taxa de analfabetismo", o que indicou que, quanto menor o nível de renda *per capita* municipal em 1980, maior a quantidade de analfabetos com mais de 15 anos de idade, no município.

Destaca-se também a correlação relativamente alta entre a renda *per capita* e o percentual da população urbana, que foi de 0,69.[5] Em contraste com o resultado obtido por Glaeser et al. (1995), tanto o crescimento populacional quanto a renda *per capita* estiveram positivamente correlacionados com o percentual da participação do setor industrial no PIB total, ambos com o valor aproximado de 0,56. Ademais, a correlação existente entre o percentual do emprego urbano e o percentual da participação do setor industrial – que assumiu o valor de 0,67 – corrobora hipóteses tais como aquelas levantadas por Pred (1966), que observou a existência de uma interdependência dos processos de urbanização e industrialização para a economia norte-americana, no século XIX.

De fato, parece óbvio que variáveis como Escolaridade e Taxa de analfabetismo sejam altamente correlacionadas. O valor assumido para essa correlação foi -0,84. O mesmo ocorre para as variáveis Mortalidade infantil e Esperança de vida ao nascer, que assumiu o valor de -0,99.

Tabela 2 – Matriz de correlação bruta das variáveis municipais[6]

	CR80-00	ESC	ANALF	ILUM	MORT	ESPVID	HOMIC	RENPC	%IND	%URB	DISCAP
CR80-00	1	0,38	-0,37	0,36	0,24	-0,23	0,19	0,40	0,60	0,52	-0,62
ESC	0,38	1	0,84	0,63	0,01	-0,01	0,10	0,82	0,58	0,75	-0,39
ANALF	-0,37	-0,84	1	-0,65	0,08	-0,08	-0,03	-0,72	-0,53	-0.66	0,33
ILUM	0,36	0,63	-0,65	1	-0,13	0,13	0,01	0,64	0,48	0,68	-0,18
MORT	0,24	0,01	0,08	-0,13	1	-0,99	0,14	-0,04	0,26	0,10	-0,36
ESPVID	-0,23	-0,01	-0,08	0,13	-0,99	1	-0,13	0,04	-0,25	-0,10	0,36
HOMIC	0,19	0,10	-0,03	0,01	0,14	-0,13	1	0,10	0,17	0,17	-0,14
RENPC	0,40	0,82	-0,72	0,64	-0,04	0,04	0,10	1	0,56	0,69	-0,40
%IND	0,60	0,58	-0,53	0,48	0,26	-0,25	0,17	0,56	1	0,67	-0,61
%URB	0,52	0,75	-0,66	0,68	0,10	-0,10	0,17	0,69	0,67	1	-0,35
DISCAP	-0,62	-0,39	0,33	-0,18	-0,36	0,36	-0,14	-0,40	-0,61	-0,35	1

Fonte: Elaborado pelo autor.

5 Tal correlação foi bastante discutida pelos autores da *Urban Economics*, a começar por Jacobs (1969), e remete a uma maior produtividade do trabalho advinda de economias de aglomeração em centros urbanos, e que, em alguma medida, traduz-se no crescimento da remuneração dos trabalhadores [Sobre essa discussão, ver, por exemplo, Galinari (2006) e Fingleton (2003)].

6 Para descrição das abreviações, ver Apêndice 1.

A distribuição geográfica da população do Estado para os anos de 1980 e 2000 é ilustrada nas Figuras 9 e 10. Ambas as figuras mostram, claramente, a concentração da população na parte leste do Estado e a manutenção desse padrão entre o período analisado.
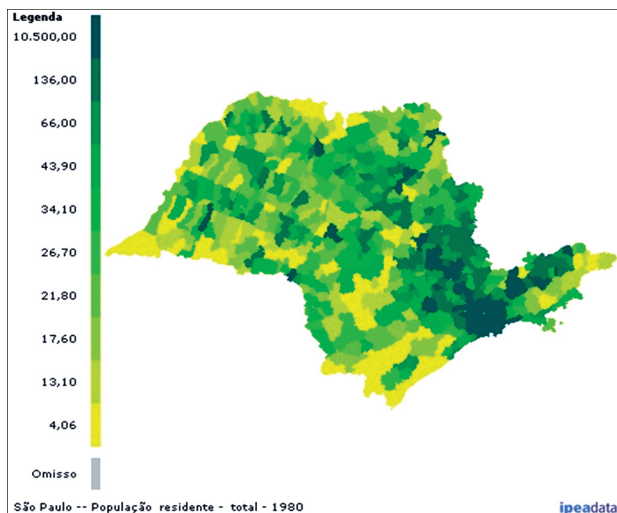
Figura 9. Distribuição espacial da população paulista em 1980.
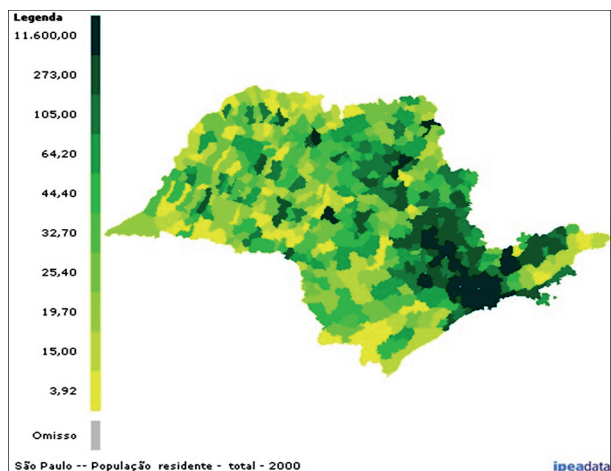
Figura 10. Distribuição espacial da população paulista em 2000.

Dado que as Figuras 9 e 10 não mostram, de forma clara, o dinamismo do comportamento demográfico no período, parte-se para uma análise exploratória de dados espaciais no intuito de entender melhor o processo.

A Análise Exploratória de Dados Espaciais (AEDE) auxilia na obtenção de evidências mais consistentes sobre a existência ou não de uma concentração geográfica da taxa de crescimento econômico no Estado de São Paulo. Através da Figura 11, é possível visualizar que a RMSP e seu entorno, com exceção da cidade de São Paulo propriamente dita, consiste em um regime espacial[7] importante, formando um polo de municípios com crescimento elevado em relação às demais regiões do Estado.[8] Uma outra região de destaque é aquela formada pela parte oeste do estado, com uma taxa de variação negativa na maioria dos municípios pertencentes a essa região. Além disso, nota-se que a região central se mantém com uma taxa de crescimento intermediária à apresentada pela RMSP e seu entorno e àquela aparentemente apresentada pela região oeste do Estado. A presença de tais polos de crescimento e de estagnação reforça o pressuposto inicial de existência de fatores espaciais influenciando a taxa de crescimento dos municípios paulistas.

A presença de *clusters* espaciais de crescimento e estagnação pode ser confirmada pelos resultados fornecidos pelo instrumental LISA (*Local Indicators of Spatial Association*). A metodologia LISA possibilita uma análise local do padrão espacial apresentado pelos dados, e leva em consideração a influência espacial em determinadas regiões, enquanto outras regiões não apresentam agrupamentos estatisticamente significantes.

---

7 Conforme Abreu et al. (2005), utiliza-se do conceito de regime espacial com referência a modelos nos quais a amostra é dividida em grupos, de acordo com os valores tomados por uma variável com dimensão espacial, por exemplo, Norte e Sul (conforme a latitude), ou tropical, subtropical e temperado (conforme o clima da região).

8 Destaca-se que a distorção demográfica mencionada anteriormente foi corrigida tanto para a construção do mapa quanto para as demais análises.
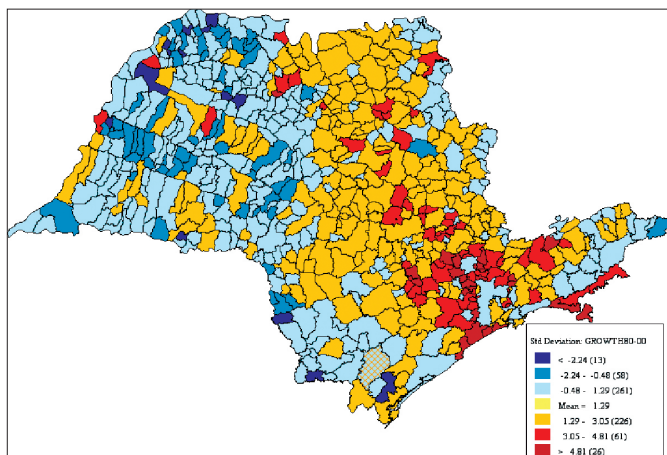
Figura 11. Distribuição espacial das taxas de crescimento anuais dos municípios paulistas (média de 1980 a 2000).

O mapeamento dos resultados obtidos para os municípios paulistas é ilustrado nas Figuras 12 e 13, que corroboram os resultados apresentados na Figura 11 e indicam a existência de duas áreas de concentração claramente distintas. A Figura 12 apresenta o gráfico de dispersão de Moran, que concentra a maior parte dos dados no 1º e 3º quadrantes, confirmando a presença de algum grau de associação espacial para uma matriz de pesos do tipo rainha. O 1º quadrante refere-se aos municípios com padrão alto-alto de crescimento, e o 3º quadrante, aos municípios com padrão baixo-baixo. O cálculo da estatística I de Moran para uma matriz de contiguidade binária "rainha" de primeira ordem assumiu o valor de 0,4591.

Por sua vez, a Figura 13 auxilia na localização dos *clusters* identificados pela Figuras 11 e 12. A região que apresenta o primeiro padrão, do tipo alto-alto, situa-se, basicamente, na região leste do Estado. A segunda área de concentração está localizada na porção noroeste-oeste e apresenta um padrão do tipo baixo-baixo. Tal organização remete à existência de uma aglomeração de municípios com baixos níveis de crescimento, cercados por municípios que também apresentaram baixo crescimento.
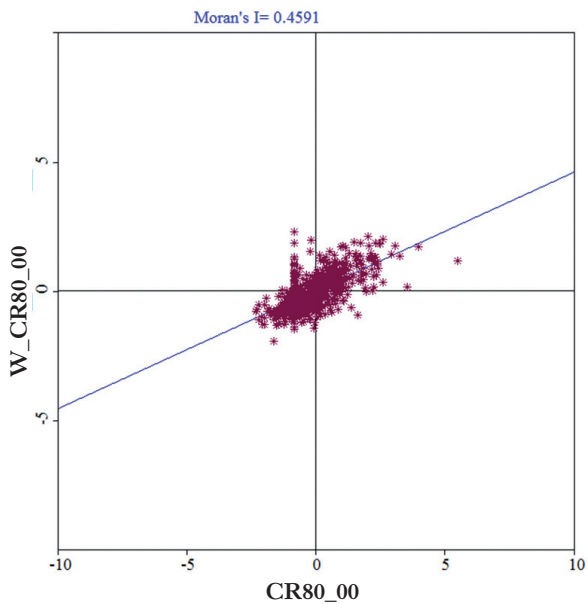
Figura 12. Gráfico de dispersão de Moran.
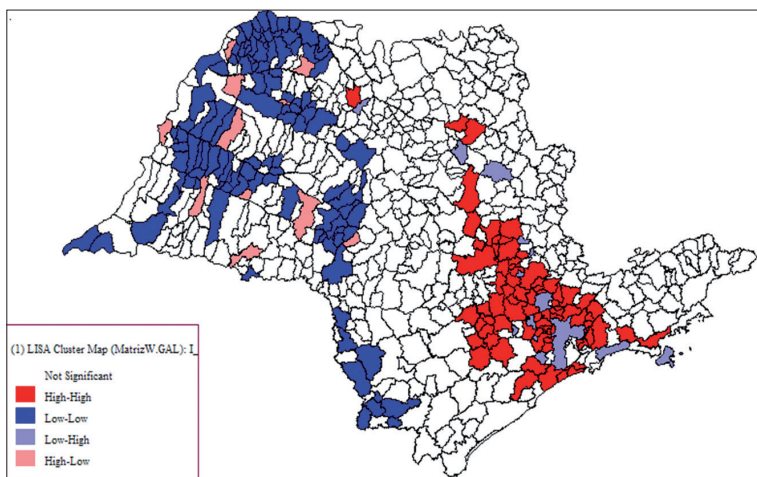
Figura 13. Mapeamento dos resultados obtidos pela metodologia LISA.

Dada a concentração geográfica verificada para a taxa de crescimento dos municípios paulistas, a estratégia adotada, a partir de então, consiste em estimar as equações de crescimento combinadas com os modelos econométricos espaciais. Busca-se, assim, identificar o modelo mais adequado ao padrão espacial apresentado pelos dados. Além disso, diferentes especificações da matriz W foram utilizadas como forma de testar a robustez dos resultados do modelo.

4
Resultados Econométricos

Uma vez identificada a existência de padrões espaciais de crescimento, denota-se a necessidade de incluir, no modelo, variáveis que captem e quantifiquem esse tipo de influência. Nesse sentido, além das variáveis municipais individuais relacionadas ao crescimento da produtividade e qualidade de vida, como proposto por Glaeser et al. (1995), o modelo considera os efeitos das externalidades espaciais. Estas podem se manifestar de dois modos: primeiro, por meio da defasagem espacial, posto que, à medida que uma cidade cresce, pressupõe-se que esta deva influenciar o crescimento de seus vizinhos. E, segundo, a influência espacial pode ser derivada de variáveis omitidas que se manifestam por meio da autocorrelação dos resíduos. Operacionalmente, o primeiro refere-se ao modelo espacial autorregressivo (SAR), e o segundo refere-se ao modelo de erro espacial (SEM).

Acrescenta-se, ainda, ao conjunto de equações o modelo espacial de *Durbin*, que também será utilizado para identificar possíveis efeitos de externalidades gerados pelas variáveis explicativas. A estimação das variáveis explicativas defasadas também possui uma segunda função relevante, que é sua utilização como um conjunto de variáveis de controle, o que possibilitou maior robustez ao modelo original.

Em resumo, o modelo estimado corresponde ao modelo de crescimento proposto por Glaeser et al., acrescido dos respectivos parâmetros espaciais. Formalmente, o modelo completo, ou seja, com a presença de todos os possíveis efeitos espaciais é o seguinte:

$$CR80_00 = \rho W_1 CR80_00 + X'_{i,1980}\beta + \theta W_2 X'_{i,1980} + \mu$$

$$\mu = \lambda W_3 \mu + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n) \qquad (5.a)$$

em que *CR80-00* é um vetor (571×1) com o percentual da taxa de crescimento anual de todos os municípios paulistas; $X'_{i,1980}$ é uma matriz (571×k) contendo o conjunto das variáveis explicativas mais a coluna de 1's correspondente ao termo de intercepto, μ é um vetor com os termos aleatórios e ε corresponde ao termo de erro não correlacionado. W_1, W_2 e W_3 são matrizes de contiguidade normalizadas pelas linhas, que, a princípio, não estão definidas formalmente. ρ, θ, λ são parâmetros.

O conjunto de variáveis explicativas do crescimento municipal foi apresentado no capítulo 3 e remete às variáveis representativas do nível de renda da população, nível educacional, infraestrutura, composição socioeconômica e distância à capital estadual, todas referentes ao período inicial, 1980.

Além disso, conforme a abordagem empregada por Ciccone & Hall (1996) e Silva Júnior (2007), incluem-se, no modelo, variáveis que busquem captar os efeitos da aglomeração. Nesse sentido, utiliza-se das variáveis "logaritmo da população", para captar os efeitos positivos, "área municipal", como controle, e a "forma quadrática do logaritmo da população", prevendo que, a partir de certo ponto, os custos de congestionamento devam superar os benefícios da aglomeração.

Frente às diversas abordagens econométricas possíveis, resolveu-se dividir esta seção em quatro partes, em que cada uma corresponde aos resultados do modelo relacionados a uma matriz de pesos espaciais específica.

Resultados do modelo com a matriz binária tradicional (rainha)

A primeira especificação para a matriz W do modelo consiste na matriz binária clássica, com seus elementos $w_{ij} = 1$, se os municípios i e j possuem fronteiras ou vértices comuns; $w_{ij} = 0$, caso contrário. A diagonal principal é composta de zeros, e a matriz **w** considerada foi normalizada ao dividir cada elemento pela soma dos elementos não nulos de sua respectiva linha.

Na Tabela 3, são apresentados os resultados do modelo econométrico por MQO e os resultados dos testes de autocorrelação espacial, bem como o modelo de Máxima Verossimilhança (MV) para a abordagem indicada pelos resultados dos testes.

A estatística I de Moran foi utilizada para a identificação de algum tipo de autocorrelação espacial, já que o teste I não apresenta contrapartida em nenhuma hipótese alternativa específica. Foram usados, também, os testes de Multiplicador de Lagrange (LM) para definir qual o tipo de autocorrelação espacial adequado ao processo gerador dos dados. Os testes para defasagem espacial, LM-LAG, e erro espacial, LM-ERR, testam a hipótese nula de $\rho = 0$ e $\lambda = 0$ na equação (5.a). Ambos os testes seguem uma distribuição χ^2 com 1 grau de liberdade. A identificação do tipo de autocorrelação espacial é realizada também com o auxílio dos testes LM robustos e dos testes Wald e Razão de Verossimilhança (LR).

Por um lado, a rejeição da hipótese nula no modelo de defasagem espacial implica que os estimadores de MQO são viesados e ineficientes; por outro, a rejeição da hipótese nula para o modelo de erro espacial indica que os estimadores de MQO são não viesados, mas não são eficientes (Anselin, 1988).

Na escolha das variáveis do modelo, a alta correlação entre duas variáveis, como apontado na Tabela 2, foi decisiva. Dessa forma, a variável Esperança de Vida ao Nascer foi excluída por possuir correlação bruta elevada, -0,99, com a variável mortalidade infantil.[1]

1 A escolha entre uma ou outra variável foi realizada de forma arbitrária.

Tabela 3 – Resultados da estimação por MQO e MV do modelo de crescimento econômico para o Estado de São Paulo – matriz rainha

Variável dependente: CR80-00				
	MQO	**SAR**	**SDM**	
			$B_{padrão}$	$B_{externalidades}$
Termo de intercepto	2,08 (2.3318)	-1,23*** (0,46)	1,12 (2,309)	–
Logaritmo da população em 1980	0,19 (0,446)	0,55*** (0,12)	0,61** (0,244)	-0,26 (0,596)
Logaritmo da população em 1980 ao quadrado	-0,03 (0,022)	-0,05*** (0,002)	-0,05*** (0,012)	0,01 (0,027)
Área municipal	0,0006*** (0,000)	0,0005*** (0,000)	0,0004** (0,000)	0,0002 (0,000)
Renda *per capita*	0,30 (0,208)	0,19 (0,189)	0,1 (0,202)	-0,11 (0,394)
Anos médios de escolaridade	-0,45*** (0,165)	-0,27* (0,149)	-0,18 (0,165)	-0,52* (0,297)
Percentual de analfabetismo	-0,03** (0,015)	-0,02* (0,014)	-0,02 (0,015)	-0,03 (0,030)
Percentual de casas com energia elétrica	0,82* (0,486)	0,67 (0,444)	1,07* (0,551)	-1,07 (0,879)
Taxa de homicídios	0,02*** (0,005)	0,01*** (0,005)	0,01** (0,005)	0,003 (0,010)
Mortalidade infantil	0,01* (0,004)	0,01*** (0,004)	0,01* (0,004)	-0,002 (0,007)
Participação do setor industrial no PIB	1,67*** (0,386)	1,47*** (0,338)	1,4*** (0,354)	0,32 (0,748)
Participação do emprego urbano	3,01*** (0,415)	2,43*** (0,367)	2,14*** (0,407)	2,08** (0,856)
Distância à capital estadual	-0,005*** (0,000)	-0,003*** (0,000)	-0,004*** (0,001)	0,002 (0,001)
Ρ	–	0,39*** (0,05)	0,36*** (0,053)	–
Λ	–	–	–	–
R^2	0,5684	0,5826	0,6006	–
R^2 ajustado	0,5592	0,5736	0,5831	–
Log Likelihood	–	-637,424	-629,486	–
Moran	0,193	–	–	–
LM-LAG	$9,5681x10^3$	–	–	–

Continua

Tabela 3 – *Continuação*

Variável dependente: CR80-00				
	MQO	**SAR**	**SDM**	
			$B_{padrão}$	$B_{externalidades}$
LM-EL	13,8573	–	–	–
Wald	386,29	–	–	–
LR	54.4930	–	–	–

Notas: (1) Valores do desvio-padrão dos parâmetros entre parênteses; (2) para as estatísticas, os parênteses contêm os respectivos p-valores;*** significativo ao nível de 1%; ** significativo a 5%; *significativo a 10%.

Como pode ser observado na Tabela 3, o modelo sugerido mostrou-se bastante representativo, e seu coeficiente de determinação, R^2, atingiu o valor de 0,5684, ou seja, 56,84% da variação na variável dependente é explicada pelas variáveis presentes no modelo. O teste para o I de Moran do modelo de MQO rejeitou a hipótese nula de nenhuma correlação espacial, exigindo, assim, a inclusão do parâmetro espacial. Os resultados dos testes LM, LM-LAG e LM-ERR mostraram-se todos significativos. Assim, seguindo a sugestão de Florax et al. (2003), foram estimados os testes LM robustos, LM-LE e LM-EL, e o maior valor para o teste foi apresentado pela estatística LM-LE, que representa a versão robusta do teste de defasagem espacial, LM-LAG. Tem-se, portanto, que o modelo SAR mostrou-se o mais adequado para o caso de uma matriz de pesos binária ponderada pelos vizinhos diretos.

Na comparação entre os modelos, o modelo espacial de Durbin mostrou-se o mais representativo, uma vez que apresentou o maior valor para o R^2 ajustado, 0,5831. Assim, ao comparar-se o modelo de Durbin com o modelo de MQO original, conclui-se que, mesmo após a inclusão das variáveis individuais e de controle, aproximadamente 2,39% da variação na variável dependente é atribuída a alguma forma de dependência espacial.[2]

2 Isso porque o R^2 ajustado do modelo de MQO é 0,5592, enquanto para a mesma regressão, incluindo os termos espaciais, o R^2 ajustado foi de 0,5831.

Uma vez identificado o modelo mais apropriado ao padrão apresentado pelos dados, conforme uma matriz de pesos binária, a etapa seguinte consiste na análise dos resultados.

O modelo SDM estimado mostrou que, quanto maior a população do município em 1980, mais o município tendeu a crescer, o que pode ser identificado por meio do parâmetro positivo e significativo para a variável logaritmo da população em 1980. Porém, essa influência positiva tende a atingir um ponto de saturação, pois, a partir de certo nível, o tamanho do município passa a ter influência negativa em função dos efeitos de congestionamento, o que é indicado pelo sinal negativo da forma quadrática do logaritmo da população. Esses resultados mostram-se ainda mais robustos à medida que a variável de controle, área municipal, também apresenta valor positivo e significante.

A interpretação de alguns parâmetros do modelo exige um certo nível de cautela. Os indicadores da quantidade de homicídios e da taxa de mortalidade infantil também se mostram positivos ao crescimento. Porém, todo e qualquer modelo econométrico deve ser analisado com os devidos cuidados, e o modelo por si só não diz tudo. Tal fato pode ser facilmente entendido, uma vez que as regiões de maior aglomeração de pessoas do Estado são acompanhadas por indicadores mais elevados de número de homicídios e mortalidade infantil e, não propriamente, que a quantidade de homicídios e mortalidade infantil provocam crescimento. Consiste em um problema de endogeneidade que não foi discutido neste trabalho.

A variável indicativa para a distância à capital estadual mostrou-se estatisticamente significativa, mostrando que quanto mais próximo da capital estadual mais o município tendeu a crescer, ou vice-versa.

Em discordância com os resultados de Glaeser et al. (1995) para os municípios dos EUA; no caso do Estado de São Paulo, a participação do PIB do setor industrial no período inicial também se mostrou significante, evidenciando a importância do setor para o crescimento dos municípios paulistas. A participação do emprego urbano também foi altamente significativa, ao indicar que, tudo o mais constante, aquelas cidades que tinham um percentual maior de trabalhadores na zona urbana, no ano de 1980, tiveram uma tendência maior ao

crescimento, o que fornece evidência adicional a favor da importância da aglomeração no processo de crescimento das cidades.

Os resultados obtidos para a participação da indústria na economia municipal pode ser um indicativo do diferente estágio de desenvolvimento vivido pela economia paulista. No modelo estimado, um município mais industrializado possui um potencial de crescimento maior, o que destoa dos resultados alcançados por Glaeser et al. (1995) para os municípios norte-americanos.

Um *survey* sobre o papel do nível educacional também pode ser extraído do conjunto de modelos estimados. Nos resultados da estimação do modelo SAR, a variável representativa dos anos médios de estudo apresentou sinal negativo e significante, indicando que, quanto maior o nível educacional inicial da população, menos o município tendeu a crescer.

Como uma primeira hipótese para a explicação desse fato, pode-se supor que a atração de pessoas menos educadas de outras regiões foi um fator determinante para o processo de crescimento do município. Entretanto, o sinal negativo e significante para a variável analfabetismo produz evidências contrárias a essa interpretação.

Por sua vez, uma segunda hipótese seria a de que um maior nível educacional da população municipal impulsionou a expulsão dos cidadãos daquelas cidades menos favorecidas de oportunidades de trabalho em direção à capital e aos centros regionais. Nessa linha, os resultados do modelo SDM favorecem essa interpretação, pois aponta para a presença de externalidades negativas para a variável anos de estudo. A escolaridade média dos municípios não se mostra significativa ao crescimento, mas o nível de escolaridade dos municípios vizinhos se apresenta como um fator negativo. Assim, aquele município que possuía um vizinho com nível educacional elevado esteve mais sujeito a perder população, o que pode ser uma evidência do alto poder de atração que o município com alta escolaridade tem sobre a população educada de seus vizinhos.

E por último, uma terceira hipótese seria a de que a variável escolaridade média de pessoas com 25 anos ou mais não consiste em uma variável representativa do nível de capital humano de uma

região. Uma afirmação que precisa ser tratada com maior afinco, o que não foi realizado neste trabalho.

Destaca-se ainda o elevado valor encontrado para o parâmetro de defasagem espacial, ρ, que para o modelo SDM, foi de 0,36 e apresentou elevada significância estatística. Esse fato corrobora a hipótese de algum tipo de externalidade espacial atuando sobre as taxas de crescimento dos municípios paulistas. Além disso, torna os resultados do modelo ainda mais robustos.

Uma segunda variável defasada importante, no modelo SDM, foi a participação do emprego urbano, que se mostrou significante e positiva, reforçando o poder de influência entre os municípios, ou seja, um município que apresenta um elevado percentual do emprego urbano tende a influenciar, positivamente, o crescimento dos municípios vizinhos.

Por fim, ao estimar-se o modelo espacial de Durbin, a variável representativa da infraestrutura municipal, percentual de domicílios com energia elétrica, passa a ser significativa, um indicador da relevância da qualidade da infraestrutura municipal para o crescimento do município.

Alternativamente à matriz binária tradicional, outras formas de especificação para a matriz W foram testadas. Assim, nos tópicos que se seguem, são apresentados os resultados para o modelo com novas especificações para a matriz de pesos espaciais.

A matriz de distância geográfica

O modelo da Tabela 4 é exatamente o mesmo da tabela anterior, exceto pela especificação da matriz de pesos espaciais utilizada. A matriz W, subscrita no modelo, a seguir foi construída com base em uma quantidade, *k*, fixa de vizinhos. Porém, a montagem da matriz não foi realizada de forma binária, como é o caso de Erthur & Gallo (2003), mas sim pelo inverso da distância ao quadrado entre os centroides dos *k* vizinhos mais próximos. Dessa forma, a matriz tem como base

a distância geográfica entre os municípios, porém não apresenta os problemas metodológicos apontados pelos referidos autores.[3]

O método de escolha da quantidade ótima de vizinhos seguiu a abordagem empregada em Baumont (2004), com a aplicação da estatística I de Moran sobre os resíduos da regressão de MQO para matrizes com diferentes números de vizinhos. Conforme sugestão de Baumont, foi escolhida a matriz com o maior valor para a estatística I. Para a determinação da matriz adequada, a quantidade fixa de vizinhos testada variou entre 1 e 15, sendo que aquela com 7 vizinhos alcançou o valor mais significativo para o I de Moran, 0,0348. A Tabela 4 repete os valores dos parâmetros da estimação por MQO e apresenta os resultados dos testes e da estimação por MV para a matriz de distância geográfica.

Tabela 4 – Resultados da estimação por MQO e MV para uma matriz com um número fixo de vizinhos

Variável dependente: CR80-00				
	MQO	**SAR**	**SDM**	
			$B_{padrão}$	$B_{externalidades}$
Constante	2,08 (2.3318)	1,8313 (11,2662)	2,9769*** (0,2665)	–
Logaritmo da população em 1980	0,19 (0,446)	0,1897 (2,1096)	0,2267 (0,4247)	0,0860 (0,3010)
Logaritmo da população em 1980 ao quadrado	-0,03 (0,022)	-0,0345 (0,1041)	-0,0369* (0,0207)	-0,0239 (0,0146)
Área municipal	0,0006*** (0,000)	0,0006* (0,0003)	0,0006*** (0,0002)	0,0008*** (0,0003)
Renda *per capita*	0,30 (0,208)	0,3022 (0,2922)	0,2646 (0,2032)	0,0564 (0,4605)
Anos médios de escolaridade	-0,45*** (0,165)	-0,4334* (0,2236)	-0,4086*** (0,1620)	-0,0873 (0,3403)
Percentual de analfabetismo	-0,03** (0,015)	-0,0322* (0,0191)	-0,0325** (0,0152)	-0,0560* (0,0324)

Continua

3 Ertur & Gallo (2003) alertam para os problemas metodológicos que surgem na montagem da matriz de pesos baseada no distância geográfica sem que seja considerada uma quantidade fixa de vizinhos.

Tabela 4 – *Continuação*

Variável dependente: CR80-00				
	MQO	SAR	SDM	
			$B_{padrão}$	$B_{externalidades}$
Mortalidade infantil	0,01* (0,004)	0,0081 (0,0055)	0,0080* (0,0043)	0,0115 (0,0091)
Participação do setor industrial no PIB	1,67*** (0,386)	1,6725*** (0,5659)	1,7157*** (0,3789)	0,1181 (0,8137)
Participação do emprego urbano	3,01*** (0,415)	2,9864*** (0,6550)	2,9472*** (0,4065)	0,4065 (0,8843)
Distância à capital estadual	-0,005*** (0,000)	-0,0049*** (0,0004)	-0,0049*** (0,0004)	0,0003*** (0,0008)
ρ	–	0,0780 (0,0484)	0,0510 (0,0141)	–
λ	–	–	–	–
R^2	0,5684	0,5695	0,5801	–
R^2 ajustado	0,5592	0,5602	0,5617	–
Log Likelihood	–	-667,6414	-661,1135	–
Moran	0,0348	–	–	–
LM-LAG	4,3347 (0,0373)	–	–	–
LM-ERR	1,6567 (0.198)	–	–	–
Wald	1,2001 (0,2733)	–	–	–
LR	1,6833 (0,1945)	–	–	–

Notas: (1) Valores do desvio-padrão dos parâmetros entre parênteses; (2) para as estatísticas, os parênteses contêm os respectivos p-valores;*** significativo ao nível de 1%; ** significativo a 5%; *significativo a 10%.

Para o caso da matriz de pesos construída por meio da distância entre os centroides de um número fixo de vizinhos, os testes de autocorrelação sobre os resíduos da regressão de MQO não indicam a presença de autocorrelação espacial, exceto para o caso do teste LM para defasagem. Os testes de Moran, LM-ERR, Wald e LR não apontaram presença de autocorrelação espacial no modelo.

A estratégia adotada foi seguir a taxonomia sugerida por Florax et al. (2003). Dado que o teste LM apresentou-se significante para

a defasagem espacial e não significante para o termo de erro, segue-se para a estimação do modelo SAR. O valor do R^2 ajustado aponta que o modelo SAR fornece uma pequena contribuição em relação ao modelo de MQO. O parâmetro para a defasagem espacial, ρ, apresentou significância estatística ao nível de aproximadamente 5%, assumindo o valor de 0,078.

Com o emprego da matriz de distância geográfica, alguns parâmetros do modelo sofrem alguma modificação em comparação com aqueles da matriz de conectividade binária. Em primeiro lugar, o conjunto de parâmetros relativos à influência do tamanho da população não se mostrou significante para o caso do modelo SAR. No modelo SDM, a forma quadrática do logaritmo da população, isto é, o parâmetro representativo dos efeitos de congestionamento, mostrou-se significativo a 10% de significância.

Em segundo lugar, os parâmetros representativos do nível educacional mantiveram-se significativos e inversamente correlacionados com o crescimento municipal. A variável nível de escolaridade da população manteve-se negativamente correlacionada com o crescimento municipal, bem como o grau de analfabetismo da população. O percentual de analfabetos dos municípios vizinhos também foi negativo e inversamente correlacionado com o nível de crescimento. Ao contrário do modelo anterior, este não apontou para a influência da escolaridade média dos municípios vizinhos.

Os parâmetros relativos ao papel da participação da indústria e do emprego urbano mantiveram-se altamente significativos e com valores elevados. A distância à capital estadual manteve sua relação inversa com o crescimento municipal, pois quanto mais distante da capital, *ceteris paribus*, menor tende a ser a taxa de crescimento média do município.

Por fim, no modelo de Durbin, a variável representativa da infraestrutura municipal também passa a apresentar-se estatisticamente significativa e positivamente correlacionada com o crescimento municipal.

Em resumo, a especificação para a matriz de pesos formada pela distância geográfica apresentou algumas alterações em relação à

interação espacial definida pela matriz de contiguidade de 1º ordem, porém enfatiza os efeitos da participação da indústria e do emprego urbano, além da importância da distância à capital do Estado e do papel da infraestrutura no crescimento. Adicionalmente, os resultados para o modelo SDM reforçam a relevância em estimar-se o modelo espacial controlando para os efeitos das defasagens das variáveis explicativas. A importância desse tipo de controle ficou ainda mais evidente no caso desta subseção.

A matriz de distância econômica

No sentido de contribuir com a discussão sobre o emprego da matriz de contiguidade mais adequada, procurou-se, em linha com Case & Rosen (1993) e Conley & Ligon (2002), adotar uma medida para a matriz W que leva em consideração uma medida de "distância econômica" entre as unidades espaciais. Por distância econômica entende-se a similaridade entre a composição da economia dos municípios, na qual a noção de proximidade perde importância e dá espaço para a ideia de semelhança, ou seja, os municípios mais parecidos possuem maior poder de influência uns sobre os outros.

Na prática, a matriz testada calcula as distâncias como diferenças de valores para um mesmo indicador entre duas localidades. Este trabalho classifica a economia municipal em três setores, mais especificamente: agricultura, indústria e serviços. A decomposição pode avançar à medida que se dispõe de dados para tal.

A construção da matriz de pesos "econômica" deu-se da seguinte forma: suponha a divisão do PIB de cada município em N setores diferentes, tal que a posição de cada município seja um conjunto de coordenadas no espaço Θ^n, a proporção de cada setor de atividade no PIB total. Dessa forma, o elemento w_{ij} da matriz W seria então o inverso da distância entre a composição de cada setor i e j neste Θ^n (zero, se i=j). Considera-se, portanto, um espaço com uma divisão simples, com distinção entre apenas três setores: indústria, agricultura e serviços, o qual foi usado para montar a matriz W normalizada

pelas linhas. O princípio basilar é que os efeitos de transbordamento decorrem da dinâmica de crescimento da produtividade setorial. Os resultados são mostrados na Tabela 5.

Tabela 5 – Resultados da estimação por MQO e MV com distância econômica na matriz W

Variável dependente: CR80-00				
	MQO	SAR	SDM	
			$B_{padrão}$	$B_{externalidades}$
Constante	2,08 (2.3318)	-1,1523*** (0,4258)	1,5020 (1,9169)	–
Logaritmo da população em 1980	0,19 (0,446)	0,5357*** (0,1241)	0,5934** (0,3020)	-0.3144 (0,5119)
Logaritmo da população em 1980 ao quadrado	-0,03 (0,022)	-0,0488*** (0,0026)	-0,0520*** (0,0152)	0.0168 (0,0233)
Área municipal	0,0006*** (0,000)	0,0005*** (0,0001)	0,0004*** (0,0002)	0.0002 (0,0003)
Renda *per capita*	0,30 (0,208)	0,1962 (0,1892)	0,0918 (0,2021)	-0.0235 (0,3876)
Anos médios de escolaridade	-0,45*** (0,165)	-0,2744* (0,1494)	-0,1853 (0,1650)	-0.5201* (0,2961)
Percentual de analfabetismo	-0,03** (0,015)	-0,0246* (0,0140)	-0,0215 (0,0147)	-0.0315 (0,0296)
Percentual de casas com energia elétrica	0,82* (0,486)	0,6634 (0,4441)	1,0923** (0,5516)	-1.1839 (0,8752)
Taxa de homicídios	0,02*** (0,005)	0,0149*** (0,0047)	0,0117** (0,0048)	0.0015 (0,0100)
Mortalidade infantil	0,01* (0,004)	0,0094*** (0,0040)	0,0075* (0,0044)	-0.0001 (0,0074)
Participação do setor industrial no PIB	1,67*** (0,386)	1,4668*** (0,3384)	1,3970*** (0,3557)	0.4936 (0,7396)
Participação do emprego urbano	3,01*** (0,415)	2,4387*** (0,3675)	2,1681*** (0,4099)	1.9856*** (0,8470)
Distância à capital estadual	-0,005*** (0,000)	-0,0029*** (0,0002)	-0,0045*** (0,0010)	0.0020* (0,0011)
ρ	–	0,3860 (0,0541)	0,3540 (0,0524	–
λ	–	–	–	–
R^2	0,5684	0,5826	0,6007	–

Continua

R^2 ajustado	0,5592	0,5736	0,5832	–
Log Likelihood	–	-637,4356	-629,6327	–
Moran	0,1951	–	–	–
LM-LAG	$2{,}4612*10^3$	–	–	–
LM-ERR	53,5077	–	–	–
LM-LE	20,5539	–	–	–
LM-EL	12,2540	–	–	–
Wald	350,954	–	–	–
LR	54,0210	–	–	–

Notas: (1) Valores do desvio-padrão dos parâmetros entre parênteses; (2) para as estatísticas, os parênteses contêm os respectivos p-valores;*** significativo ao nível de 1%; ** significativo a 5%; *significativo a 10%.

Os testes para o modelo de MQO utilizando a matriz de distância econômica apontam para a existência de autocorrelação espacial no modelo. Como ambos os testes LM mostraram-se significativos, os testes robustos foram utilizados, conforme sugerido por Florax et al. (2003). O resultado dos testes robustos aponta para a existência de defasagem espacial no modelo, exigindo a estimação do modelo autorregressivo espacial SAR.

Os resultados do modelo SAR indicam uma forte influência do parâmetro espacial, que foi altamente significativo, e assumiu o valor de 0,386. O alto valor para ρ ratifica o poder de influência de municípios com economias similares, o que valida a tentativa de considerar-se a composição socioeconômica para captar efeitos de externalidades que transbordam as fronteiras geográficas.

Os resultados para os parâmetros do modelo são similares aos modelos anteriormente considerados, principalmente, para o caso do modelo com a matriz binária. Destaca-se, novamente, a significância estatística do conjunto de variáveis que representam a influência do tamanho do município sobre as taxas de crescimento, tanto para o log da população quanto para sua forma quadrática, o que corrobora a hipótese dos efeitos positivos do tamanho populacional até um determinado patamar. As variáveis que representam a participação do setor industrial e do emprego urbano mantiveram-se estatisticamente significantes.

Ao estimar o modelo de Durbin, observou-se uma melhora considerável da representatividade do modelo, sendo que o valor do R^2 atinge o valor de 0,6007, com o respectivo R^2 ajustado com valor de 0,5832. Os parâmetros do modelo SAR continuaram significativos. Além disso, dois parâmetros indicativos dos *spillovers*, nas variáveis explicativas, merecem destaque no modelo SDM.

O primeiro refere-se à variável representativa do nível de escolaridade do município "vizinho" (parecido), que volta a ser significativa e inversamente relacionada à taxa de crescimento municipal. Para o caso da matriz de distância econômica, a tese do alto poder de atratividade das cidades "vizinhas" reflete o poder de atração de municípios com características econômicas similares.

O segundo parâmetro, relacionado à variável para a participação do emprego urbano da "vizinhança", passa a ser significativo, indicando que, *ceteris paribus*, uma elevada participação do emprego na zona urbana do município "vizinho" influencia, positivamente, a taxa de crescimento municipal. Desse modo, mais uma vez, a estimação da defasagem das variáveis explicativas mostrou-se importante para o correto tratamento e interpretação dos resultados obtidos.

A matriz hierárquica

Os resultados dos modelos anteriores apontam para o modelo de defasagem espacial como definidor da interação espacial dos dados, independentemente do tipo de especificação para a matriz W. Portanto, segundo os resultados obtidos, a taxa de crescimento de um determinado município depende não apenas de seus próprios fatores, mas também dos fatores presentes em sua vizinhança.

No primeiro modelo, a distância entre dois municípios foi medida por fronteiras diretas ou vértices comuns. No segundo, a estratégia foi a utilização do inverso da distância ao quadrado com um número preestabelecido de vizinhos. No terceiro, o critério empregado foi a semelhança entre a economia dos municípios, sendo que os municípios mais parecidos apresentavam poder de influência mútua maior.

Em Abreu et al. (2005), os autores alertam para o fato de que a escolha da matriz de pesos adequada não deve seguir uma regra preestabelecida, mas sim um modelo teórico factível de ser testado. Os autores criticam os métodos da econometria espacial padrão de identificação e estimação dos modelos espaciais e defendem a tese de que a escolha da matriz de pesos deve ter uma proximidade maior com a teoria.

Com a finalidade de aproximar este estudo das ideias fornecidas pelos autores supracitados, e tomando como base a análise exploratória de dados espaciais e os resultados dos modelos anteriores, este trabalho propõe uma matriz de pesos específica para o caso dos municípios paulistas.

Por um lado, a distância geográfica parece ser um fator determinante das taxas de crescimento dos municípios. De certo modo, qualquer município pertence à vizinhança de qualquer outro, dependendo do critério adotado. Contudo, a importância relativa de cada município em uma vizinhança particular varia inversamente com a distância. Nesse caso, a similaridade com o modelo autorregressivo das séries de tempo é óbvia, uma vez que, quanto maior a distância entre os municípios, menor sua interação potencial.

Por outro lado, parece razoável supor que uma grande cidade é provavelmente menos afetada pelo que acontece em cidades próximas, do que um município pequeno. Dessa forma, seguindo Moreno & Trehan (1997), resolveu-se ajustar a ponderação para o tamanho do município, com a criação de uma matriz W que considera tanto o efeito inversamente proporcional da distância, quanto o efeito direto do tamanho do município vizinho. O ajuste foi feito ao multiplicar os pesos do inverso das distâncias pelos logaritmos do tamanho da população municipal. O princípio-base para a construção da matriz W a que se denomina *hierárquica*[4] demonstra que é preferível estar próximo a uma grande economia a uma economia de pequeno porte.

---

4 Denominou-se a matriz de pesos *hierárquica* porque determina um ordenamento aos municípios, ponderando-os de forma a atribuir maior relevância aos municípios maiores.

Operacionalmente, a matriz W hierárquica foi montada da seguinte maneira:

$$w^*_{ij} = \frac{\log(população_i)}{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2}, \text{ se } i \neq j$$

$$w^*_{ij} = 0, \text{ se } i = j \qquad (5.5.1)$$

Sendo que a notação permanece a mesma das anteriores.

A distância limítrofe escolhida diz respeito a distância mínima que faz que cada um dos municípios tenha, no mínimo, um vizinho. O resultado do cálculo dessa distância d_{ij} mínima foi de aproximadamente 61 quilômetros. Dado que a área dos municípios do Estado varia de forma intensa, a escolha de uma distância limite gera regimes espaciais variados em termos de quantidade de municípios. Ertur & Gallo (2003) chamam atenção para os problemas metodológicos advindos desse tipo de abordagem. Os resultados dos testes para a matriz hierárquica são apresentados na Tabela a seguir.

Tabela 6 – Resultados dos testes de autocorrelação espacial com o emprego da matriz de pesos hierárquica

Moran	0,0226 (0,2165)
LM-LAG	2,6659 (0,1025)
LM-ERR	1,2595 (0,2618)
Wald	0,8282 (0,3628)
LR	1,2668 (0,2604)

Nota: p-valores entre parênteses.

O modelo estimado que fornece as estatísticas apresentadas na Tabela 6 possui a mesma estrutura dos modelos anteriores e difere-se apenas no emprego da matriz W hierárquica para a realização dos testes de autocorrelação espacial. Os resultados dos testes para a matriz hierárquica não apontaram para a presença de autocorrela-

ção espacial para o conjunto dos municípios do Estado, o que torna preferível, assim, o modelo de MQO original.

O modelo de MQO já foi apresentado anteriormente, sendo que os resultados obtidos, nesse caso, são exatamente os mesmos dos modelos anteriores.

Em suma, a tentativa de serem substituídas as matrizes W tradicionais por uma matriz de pesos específica não produziu resultados concretos. No entanto, com os resultados obtidos, não se pode rejeitar, em definitivo, a funcionalidade de tal estratégia, uma vez que parte dos problemas metodológicos pode ser atribuída à não definição de um número fixo de vizinhos, conforme colocação de Ertur & Gallo (2003). Em outras palavras, a utilização do inverso da distância ao quadrado sem a definição de um número fixo de vizinhos pode ter sido a origem da inadequação da matriz hierárquica às unidades espaciais. Além disso, o padrão teórico proposto pode não ser o mais adequado ao processo gerador dos dados sobre os municípios de São Paulo, o que abre espaço para pesquisas futuras.

Considerações finais

Este estudo aborda temas de crescimento econômico e externalidades espaciais tendo em vista sua relevância na teoria econômica. Os resultados possibilitaram a investigação sobre quais variáveis são correlacionadas com as taxas de crescimento municipal no Estado de São Paulo, fornecendo assim, uma base para indicações de políticas públicas de estímulo ao crescimento.

A despeito dos desafios teóricos e empíricos comumente enfrentados pelos estudiosos do crescimento, buscou-se avançar na questão, por meio da combinação de ferramentas teóricas e econométricas, basicamente, de três grandes campos da ciência econômica *mainstrean*, a saber: crescimento endógeno, nova geografia econômica e econometria espacial.

Destacou-se o caráter dinâmico dos modelos estimados, com o emprego de variáveis do período inicial para explicar as taxas de crescimento do período posterior, 1980-2000. Adicionalmente, o trabalho investigou a presença de efeitos de transbordamento entre as variáveis municipais, ao inserir a questão espacial como determinante das taxas de crescimento dos municípios.

Em linhas gerais, pode-se dizer que o modelo construído foi satisfatório no sentido de explicar o crescimento das cidades de São Paulo.

O mapeamento realizado para o Estado permitiu identificar a presença de regimes espacias de crescimento, principalmente, nas regiões leste e oeste de São Paulo. A região leste é caracterizada por municípios de alto crescimento circundados por municípios que também o apresentam. Por sua vez, a região oeste é marcada por um *cluster* espacial de baixo crescimento, o que lhe atribui, na maioria das vezes, taxas negativas. Tais regimes foram estatisticamente comprovados por meio dos indicadores LISA.

De modo geral, tudo o mais mantido constante, as cidades mais próximas à região metropolitana de São Paulo tiveram propensão a um crescimento relativamente mais alto. O parâmetro estatisticamente significativo em todos os modelos para a variável Distância à Capital confirma essa afirmação.

O papel do tamanho do município no período inicial também foi discutido. Os resultados apontam para a existência do padrão "U" invertido como definidor do crescimento dos municípios. A princípio, o tamanho do município propende a ter influência positiva no crescimento. Entretanto, a partir de certo patamar, esse indicador passa a apresentar efeitos negativos em função das deseconomias de escala.

A maioria das variáveis indicativas da produtividade dos municípios foi significativa, e os sinais estiveram de acordo com as expectativas, exceto para a variável renda *per capita* que não se mostrou significativa em nenhum dos modelos estimados.

A variável escolaridade média também foi uma exceção porque, diferentemente de certo consenso dos modelos teóricos e empíricos fornecidos pela literatura, não apresentou correlação positiva com o crescimento econômico; ao contrário, foi negativa e estatisticamente significante.

A estimação do modelo espacial de Durbin forneceu uma constatação interessante para tal. O modelo identificou que o nível de escolaridade da vizinhança municipal influencia, de forma negativa, o crescimento, ou seja, quanto mais bem educada for a vizinhança de um município, *ceteris paribus*, menor o nível de crescimento deste. Isso pode ser um indício do poder de atratividade de pessoas com alto nível de escolaridade, que migram para os centros regionais em busca de novas oportunidades.

Os resultados das variáveis indicativas da qualidade de vida não fornecem evidências conclusivas. Para a NGE, existe uma relação inversamente proporcional entre as variáveis Taxa de Homicídios e Mortalidade Infantil com o crescimento econômico. As estatísticas dos modelos indicam uma relação positiva e significativa que, provavelmente, advém da endogeneidade existente entre as referidas variáveis e a variável dependente. Uma vez que a aglomeração de pessoas tende a ser acompanhada por indicadores elevados de mortalidade infantil e quantidade de homicídios, é mais provável a existência do fator endógeno no modelo.

A infraestrutura municipal também se mostrou relevante para a taxa de crescimento, bem como a composição da economia do município. Os resultados mostram que aquele município que apresentava maior participação da indústria em sua produção total tendeu a crescer mais. Além disso, aqueles municípios que apresentaram maior percentual de população empregada na zona urbana também tiveram, em média, maiores taxas de crescimento. Tais resultados reforçam os argumentos a favor da industrialização e urbanização do município na busca por maiores taxas de crescimento.

Os resultados das estatísticas identificaram a presença de dependência espacial no crescimento das cidades paulistas, o que permitiu quantificar os efeitos de transbordamento por meio da inclusão de um parâmetro de defasagem espacial no modelo. Quatro especificações para a matriz W foram testadas: (1) a matriz de pesos rainha, (2) a matriz de distância geográfica, (3) a matriz de distância econômica e (4) a matriz hierárquica.

O parâmetro indicativo de defasagem espacial foi positivo e altamente significativo em três das quatro abordagens. Os resultados dos testes de autocorrelação espacial foram significantes nos modelos construídos a partir da matriz binária rainha, da matriz geográfica e da matriz de distância econômica.

Os resultados para a matriz de distância econômica corroboram a tese de que municípios com as mesmas características econômicas possuem maior poder de influência mútua.

Além disso, este trabalho ressaltou a relevância em incluir o parâmetro espacial nos modelos de crescimento, bem como a inclusão

da defasagem das variáveis explicativas, no sentido de dar maior robustez aos resultados.

Todavia, a tentativa de substituir a matriz W tradicional pela matriz hierárquica não produziu resultados coerentes para o caso dos municípios paulistas. Entretanto, esse fato não significa que se deva rejeitar, em definitivo, os esforços de pesquisa nesse sentido.

A Tabela 7, a seguir, apresenta um sumário dos resultados obtidos pelos modelos estimados.

Tabela 7 – Sumário dos principais resultados obtidos nos modelos

	Matriz rainha		**Distância geográfica**		**Distância econômica**	
Variáveis	**Sinal**	**Significante**	**Sinal**	**Significante**	**Sinal**	**Significante**
Distância à capital estadual	–	Sim		Sim	–	Sim
Aglomeração	+	Sim	+	Não	+	Sim
Desaglomeração	–	Sim	–	Sim	–	Sim
Escolaridade Média	–	Não	–	Sim	–	Não
Infraestrutura	+	Sim	+	Sim	+	Sim
Part. do PIB Industrial +	Sim	+	Sim	+	Sim	
Part. do emprego urbano	+	Sim	+	Sim	+	Sim
rho	+	Sim	+	Sim	+	Sim
Defasagem escolaridade	–	Sim	–	Não	–	Sim
Defasagem Emprego Urbano	+	Sim	+	Não	+	Sim

Algumas implicações de políticas podem ser deduzidas.

Os resultados reforçam, em algum grau, os argumentos em prol da correção de desníveis educacionais e de infraestrutura entre os municípios. Adicionalmente, valida não só esforços de políticas públicas em relação a fatores que aumentem a produtividade e qualidade de vida nos municípios, como também quanto a políticas industriais, em nível municipal.

Entretanto, apesar de reforçar argumentos em favor das políticas públicas em nível individual, os modelos mostram que a distância entre os municípios é um fator crucial. Estar próximo à cidade de São Paulo, ou então, a economias mais desenvolvidas é um fator determinante para o crescimento dos municípios. Nessa linha, deve-se ressaltar a relevância de políticas públicas regionais que estimulem determinados setores da atividade econômica.

E, por fim, deve-se ressaltar a importância que a aglomeração de pessoas tem sobre as taxas de crescimento das cidades, dado que o tamanho inicial da população demonstrou ser uma variável relevante para o modelo, bem como a participação do emprego urbano. Dessa forma, o trabalho corrobora argumentos favoráveis a políticas de estímulo à aglomeração visando ao desenvolvimento regional, dado o comportamento do tipo U invertido de influência do tamanho do município no crescimento.

Apêndice
Descrição das variáveis[1]

Variação da população 1980-2000 (CR80-00). Descreve a variação do tamanho da população municipal ocorrida entre os anos de 1980 e 2000. É calculada por meio da taxa de crescimento médio anual para cada município no período considerado.

Nível de escolaridade média (ESC). Refere-se à razão entre o somatório do **número de anos de estudo completados pelas pessoas que têm 25 ou mais anos de idade e o número de pessoas nessa faixa etária.**

Taxa de analfabetismo (ANALF). Essa variável é calculada por meio do percentual da população analfabeta com mais de 15 anos de idade, relativamente à população total de cada município. Diz respeito às pessoas dessa faixa etária que não sabem ler nem escrever um bilhete simples.

Infraestrutura (ILUM). Essa variável indica o percentual de domicílios com energia elétrica em cada município.

Mortalidade infantil (MORT). Número de pessoas de cada mil nascidas vivas que não deverão completar 1 ano de vida.

Esperança de vida ao nascer (ESPVID). Expectativa de anos de vida de uma pessoa nascida no ano de referência, supondo que as

1 Informações extraídas da base de dados do Ipea – Ipeadata.

taxas de mortalidade, por idade, estimadas para anos anteriores se mantivessem constantes nos anos posteriores.

Taxa de homicídios (HOMIC). Taxa de homicídios por município (unidade).

Renda per capita (RENPC). Corresponde à razão entre o somatório da renda familiar *per capita* de todos os domicílios e o número total de domicílios no município.

Participação do setor industrial no PIB municipal (%IND). Descreve a parcela do PIB municipal referente ao setor industrial. Incluem-se no PIB Industrial, a custo de fatores, Indústrias de Transformação, Extrativa Mineral, da Construção Civil e dos Serviços Industriais de Utilidade Pública.

Participação do emprego no setor urbano (%URB). Remete à parcela do emprego correspondente ao setor urbano. Foi considerada como OCUPADA a pessoa que trabalhou nos últimos 12 meses anteriores à data de referência do Censo, ou parte deles. A pessoa que não trabalhou nos últimos 12 meses anteriores à data de referência do Censo, mas que, nos últimos 2 meses, tomou alguma providência para encontrar trabalho, foi considerada como DESOCUPADA.

Referências bibliográficas

ABREU, M. et al. Space and growth: a survey of empirical evidence and methods. *Tinbergen Institute Discussion Paper*, 129/3, 2004.

ADES, A. F.; GLAESER, E. L. Trade and circuses: explaining urban giants. *The Quarterly Journal of Economics*, v.110, n.1, p.195-227, 1995.

ALMEIDA, M. A. S. *Análise exploratória e modelo explicativo da criminalidade no Estado de São Paulo:* interação espacial (2001) Dissertação (mestrado) FCL/Unesp, Araraquara, 2007.

ALESINA, A.; RODRICK, D. Distributive politics and economic growth, *Quartely Journal of Economics,* n.109, p.465-490, 1994.

ALONSO, W. *Location and land use*: toward a general theory of land rent. Cambridge: Harvard University, 1964.

ANSELIN, L. *Spatial econometrics*: methods and models. Kluwer Academic Publishers: Dordrecht, 1988.

________. Spatial externalities, spatial multipliers and spatial econometrics. *International Regional Science Review*, v.26, n.2, p.153-66, 2003.

________.; BERA, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: ULAH A.; GILES D. E. (eds.). *Handbook of Applied Economic Statistics.* Marcel Dekker: New York, p.237-89, 1998.

________.; FLORAX, R. J. G. M. (eds.). *New directions in spatial econometrics*. Springer: Berlim, 1995.

________. (eds.). *Advances in spatial econometrics,* Springer: Berlim, 2004.

ARROW, K. (1962) Economic welfare and the allocation of resources for invention. In: LAMBERTON, D. (ed.). *Economics of Information and Knowledge.* Harmondsworth: Penguin Books, 1971.

BARRO, R. Government Spending in a Simple Model of Endogenous Growth, *Journal of Political Economy*, v.98, n.5, p.103-25, oct. 1990.

________. Economic growth in a cross-section of countries. *The Quarterly Journal of Economics,* n.2, p.407-43, 1991.

________.; SALA-I-MARTIN, X. *Economic growth*. McGraw-Hill: New York, 1995.

BAUMOL, W. J. Productivity Growth, Convergence and Welfare: what the Long Run Data Show? *American Economic Review*, v.76, n.5, p.1072-85, 1986.

BAUMONT, C. *Spatial effects in housing price models*: do housing prices capitalize urban development policies in the agglomeration of Dijon (1999)? Pôle d'Economie et de Gestion: Université de Bourgogne, 2004.

BERA, A.; YOON, M. Specification testing with locally misspecified alternatives, *Econometric Theory*, n.9, p.649-658, 1993.

BURRIDGE, P. On the Cliff-Ord Test for Spatial Autocorrelation, *Journal of the Royal Statistical Society,* v.42, p.107-8, 1980.

CANO, W. Questão regional e política econômica nacional. In: CASTRO, A. C. (org.) *Desenvolvimento em Debate*: novos rumos do desenvolvimento no mundo. Rio de Janeiro: BNDES, vol.2, p.275-307, 2002.

CASE, A.; ROSEN, H. S. Neighborhood influence and technological change. *Regional Science and Urban Economics,* v.22, p.491-508, 1993.

CASS, D. Optimum growth in an aggregative model of capital accumulation. *The Review of Economic Studies*, v.32, n.3, p.233-240, jul. 1965.

CICCONE, A. Aglomeration-effects in Europe. *European Economic Review*, v.46, n.2, feb. 2002.

________. ; HALL, R. Productivity and the density of economic activity. *American Economic Review*, v.87, n.1, mar. 1996.

CLIFF, A.; ORD, J. K. *Spatial autocorrelation*. London: Pion, 1973.

COMBES, P. P. Marshall-Arrow-Romer externalities and city growth. Paris: Ecole Nationale des Ponts-et-Cháussess, *CERAS Working Paper*, n.99-06, 2000.

CONLEY, T.G.; LIGON, E. Economic distance and cross-country spillovers, *Journal of Economic Growth*, n.7, p.157-87, 2002.

CONLEY, T.G. GMM. Estimation on cross-sectional dependence, *Journal of Econometrics*, n.92, p.1-45, 1999.

CHRISTALLER, W. (1933) *Central places in southern germany.* London: Prentice Hall, 1966.

DINIZ, C.C. (2002) A nova configuração urbano-industrial no Brasil. In: KON, A. (org.) *Unidade e Fragmentação*: a questão regional no Brasil. São Paulo: Perspectiva.

_________.; HALL, R. Productivity and the density of economic activity. *American Economic Review,* v.87, n.1, mar. 1996.; CROCCO, M.A. Reestruturação econômica e impacto regional: o novo mapa da indústria brasileira. *Nova Economia,* 6 (1), jul. 1996.

DIXIT, A. K.; STIGLITZ, J. E. Monopolistic competition and optimum product diversity. *American Economic Review,* vol.67, p.297-308, 1977.

DURBIN, J.; WATSON, G. S. Testing for serial correlation in least squares regression. *Biometrika,* v.37, p.409-28, 1950.

_________. Testing for serial correlation in least squares regression. *Biometrika,* v.38, 159-71, 1951.

ERTUR, C.; LE GALLO, J. Exploratory spatial data analysis of the distribution of regional *per capita* GDP in Europe, 1980-1985". *Papers in Regional Science RSAI,* v.82, p.175-201, 2003.

FINGLETON, B. Estimates of time to economic convergence: an analysis of regions of the European Union. *International Regional Science Review,* v.22, n.1, p.3-34, 1999.

_________. Increasing returns: evidence from local wage rates in Great Britain. *Oxford Economic Papers,* v.55, n.4, p.716-739, 2003.

FLORAX, R.; FOLMER, H. Specification and estimation of spatial linear regression models: Monte Carlo evaluation of pre-test estimators, *Regional Science and Urban Economics,* v.22, n.3, sept. 2003.

_________.; REY, S. J. The impact of misspecified spatial interaction in linear regression models. In: ANSELIN, L.; FLORAX, R. (eds.). *New Directions in Spatial Econometrics.* Springer: Berlim, 1995.

_________. Specification searches in spatial econometrics: the relevance of Hendry's methodology", *Regional Science and Urban Economics,* v.33, dec. 2003.

FUJITA, M.; THISSE, J.F. *Economics of Aglomeration*: Cities, Industrial Location and Regional Growth. Cambridge University, 2002.

FUJITA, M. et al. *The Spatial Economy: Cities, Regions and International Trade.* MIT Press: Cambridge, 1999.

GALINARI, R. *Retornos crescentes urbano-industriais e spillovers espaciais*: evidências a partir da taxa salarial no Estado de São Paulo. Dissertação (mestrado), CEDEPLAR/UFMG, Belo Horizonte, 2006.

GLAESER, E. L. et al. Growth in cities. *The Journal of Political Economy*, v.100, n.6, p.1126-52, 1992.

________. et al. Economic growth in a cross-section of cities. *Journal of Monetary Economics*, v.36, p.117-43, 1995.

GREENE, W. H. *Econometric analysis*, Prentice Hall, Nova Jersey, 1997.

GRIFFITH, D. A. Spatial autocorrelation: a primer, *resource publications in Geography*, Washington, 1987.

________. *Advanced spatial statistics*: special topics in the exploration of quantitative spatial data series. Kluwer Academic Publishers: Dordrecht, 1988.

HENDERSON, C. R. General flexibility of linear models for sire evaluation, *Journal Dairy Science*, v.57, n.8, p.963-72, 1974.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA – Ipeadata. *Dados macroeconômicos e regionais*. Disponível em: <http://www.ipeadata.gov.br>. Acesso em 6 jul. 2008.

JACOBS, J. *The economy of cities.* New York: Vintage, 1969.

KOOPMANS, T. C. On concept of optimal economic growth. In: *The economic approach to development planning*. Amsterdan: North Holland, 1965.

KNIGHT, F. Diminishing Returns to Investment. *Journal of political economy*, n.12, p.34-46, jun. 1944.

KRUGMAN, P. Increasing returns and economic geography. *Journal of political economy*, v.99, n.3, p.483-499, 1991.

________. What's new about new economic geography? *Review of Economic Policy*, v.14, n.2, p.7-17, 1998.

LEE, K. et al. Growth and convergence: a multicountry empirical analysis of the Solow growth model. *Journal of Aplied Econometrics*, v.12, p.357-392, 1997.

LE GALLO, J. & YRIGOYEN, C. C. Spatial analysis of urban growth in Spain (1900-2001). *Empirical Economics*, 2007.

LESAGE, J. P. (1999) *Spatial econometrics*. Manuscrito não publicado disponível em < http://rri.wvu.edu/regscweb.htm>.

________. A multiprocess mixture model to estimate space-time dimensions of weekly pricing of certificates of deposits. In: Anselin, L.; Florax, R. J. G. M. (eds.). *New directions in spatial econometrics*. Springer: Berlim, 359-88, 1995.

LIM, U. A spatial analysis of regional income convergence. *Planning Forum*, n.9, p.66-80, 2003.

LÖSCH, A. *The economics of location*. New Haven, Conn: Yale University, 1954.

LUCAS, R. On the mechanics economic development. *Journal of Monetary Economics,* v.12, p.3-42, 1988.

MANKIW, N. G. ________. A contribution to the empirics of economic growth. *The Quartely Journal of Economics*, v.107, n.2, p.407-37, may. 1992.

MARSHALL, A. (1920) *Princípios de economia*. São Paulo: Abril Cultural, 1982.

MAGALHÃES, A. Clubes de convergência no Brasil: uma abordagem com correção espacial. *Anais do XXIX Encontro Nacional de Economia,* Anpec, Salvador, 2001.

________. et al. Spatial dependence and regional convergence in Brazil". *Working Paper,* REAL, 00-T-11, Urbana, 2000.

MALTHUS, T. R. *An essay on the principle of population,* as it affects the future improvement of society, with remarks on the speculations of Mrs. Godwin, M. Condorcet and Others Writers. J. Johnson, V IX, London, U.K, 1798.

MONASTÉRIO, L. M.; ÁVILA, R. P. (2004) Uma análise espacial do crescimento econômico do Rio Grande do Sul. In: *Encontro Nacional de Economia*, 32, João Pessoa, PB. Anais, Belo Horizonte, Anpec. Disponível em: <http//www.anpec.org.br/encontro2004/artigos/A04A113.pdf>

MORENO, R.; TREHAN, B. Location and the growth of nations. *Journal of Economic Growth*, v.2, n.4, p.399-418, 1997.

OLIVEIRA, C. A. Externalidades espaciais e o crescimento econômico das cidades do Estado do Ceará, *Anais do X Encontro Regional de Economia do Nordeste*, Fortaleza, 2005.

PAELINCK, J.; KLAASSEN, L. *Spatial Econometrics*. Saxon House: Farnborough, 1979.

PIMENTEL, E. A.; HADDAD, E. A. Desigualdades regionais em Minas Gerais: análises espaciais do fenômeno, 1991-2000. In: *III Encontro da Associação Brasileira de Estudos Regionais*, Belo Horizonte, 2004.

PRED, A. *The spatial dynamics of economic development*. Cambridge: MIT Press, 1966.

RAMSEY, F. P. A mathematical theory of saving. *The Economic Journal*, v.38, n.152, p.543-59, dec. 1928.

RAO, C. R. On the linear combination of observations and the general theory of least squares. *Sankhyã*, Calcutta, v.7, n.3, p.237-256, 1947.

RESENDE, G. M. O crescimento econômico dos municípios mineiros: externalidades importam? *Cadernos BDMG*, n.11, p.1-96, Belo Horizonte, out. 2005.

REY, S. J.; MONTOURI, B. D. U.S. Regional income convergence: a spatial econometric perspective, *Regional Studies*, v.33, p.143-156, 1999.

RICARDO, D. (1817) *Princípios de economia política e tributação*. São Paulo: Abril Cultural, 1982.

ROMER, P. Increasing returns and long-run growth, *Journal of Political Economy*, v.94, p.1002-037, 1986.

SARTORIS, A. *Homicídios na Cidade de São Paulo*: Um Estudo de Causalidade e Autocorrelação Espaço Temporal. Tese (Doutorado), IPE/FEA/USP, São Paulo, 2000.

SCHUMPETER, J. (1911) *The theory of economic development*. London: Oxford University Press, 1934.

FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS – Seade. *Informações Socioeconômicas*. Disponível em <http//www.seade.gov.br>. Acesso em 6 jul. 2008.

SMITH, A. (1776) *A riqueza das nações*: investigação sobre sua natureza e suas causas. (Trad. port.) São Paulo: Abril Cultural, 1983.

SOLOW, R. A contribution to the theory of economic growth. *The Quartely Journal of Economics*, v.70, n.1, p.65-94, feb. 1956.

SHESHINSKI, E. Optimal accumulation with learning-by-doing. In: SHELL, K. (ed.) *Essays on the Theory of Optimal Economic Growth*. Cambridge: MIT Press, 1967.

SILVA JÚNIOR, D. *Aglomeração e desenvolvimento*: evidências para municípios brasileiros. Dissertação (mestrado), FEA/USP, São Paulo, 2007.

SILVEIRA NETO, R. M. Crescimento e *spillovers*: a localização importa? Evidências para os estados brasileiros. *Revista Econômica do Nordeste*, v.32, N.ESP, p.524-45, Fortaleza, 2001.

SWAN, T. W. (1956) Economic growth and capital accumulation. In: STIGLITZ, J. E.; UZAWA, H. (1969) *Readings in the modern theory of economic growth*. The MIT Press.

TIEFELSDORF, M.; BOOTS, B. The exact distribution of Moran's I. *Environment and Planning A*, v.27, p.985-99, 1995.

TOBLER, W. R. A computer model simulation of urban growth in the Detroit region. *Economic Geography*, n.46, p.234-240, 1970.

TYSLER, M. Econometria especial: discutindo medidas para a matriz de ponderação espacial. Dissertação (mestrado) EAE/FGV, São Paulo, 2006.

VERNER, D.; TEBALDI, E. Convergence, dynamics and geography of economic growth: the case of municipalities in Rio Grande do Norte, *World Bank Working Papper,* 3302, May. 2004. Disponível em: <http://econ.worldbank.org/view.php?type=5&id=35586>

VON THÜNEN, J. H. (1826) *The isolated state*. Oxford: Pergamon Press, 1966.

YOUNG, A. A. Increasing returns and economics progress. *The Economic Journal*, v.38, p.527-42, 1928.

SOBRE O LIVRO

Formato: 14 x 21 cm
Mancha: 23,7 x 42,5 paicas
Tipologia: Horley Old Style 10,5/14
1ª edição: 2009

EQUIPE DE REALIZAÇÃO

Coordenação Geral
Marcos Keith Takahashi